



Durham E-Theses

Statistical methods for supporting urgent care delivery

STIRLING, SARAH,GRACE

How to cite:

STIRLING, SARAH,GRACE (2011) *Statistical methods for supporting urgent care delivery*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/634/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Statistical methods for supporting urgent care service delivery

Sarah Grace Stirling

Abstract

Forecasting procedures were developed and implemented in an out-of-hours GP provider in the North East of England to ensure staffing levels were optimised, and server performance was investigated. Initial methods included linear regression to predict calls per day into a call centre, loess to predict arrival rates, and moving averages to deal with unexpected flu pandemics. We also tried to understand the behaviour of GPs and develop a fair rating system, based on their speed. Finally, we introduced some novel dissemination techniques so that the procedures could be completed by non-experts through the implementation of the RExcel software.

Declaration

The work in this thesis is based on research carried out at the Department of Mathematical Sciences at Durham University. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text. This research includes initial ideas from David Wooff, particularly in Chapter 6.

The copyright of this thesis rests with the author. No quotation from it should be published without the prior written consent and information derived from it should be acknowledged.

Acknowledgements

My thanks must go to all the staff at Northern Doctors Urgent Care, for providing an exciting case study to work on, supporting me through the process of introducing statistical methods into the company, as well as welcoming me to the North East. Particular thanks go to Ben Stobbs for helping with the data extraction, and my company supervisors John Harrison and Petrina Smith for supporting this project.

Thank you to David Wooff for the guidance and support from start to finish of this research process. His feedback and suggestions throughout every stage of this process contributed to the project being such a success.

Nomenclature

A&E	Accident and Emergency
AHT	Average Handling Time
DoH	Department of Health
HPA	Health Protection Agency
KPI	Key Performance Indicator
KTP	Knowledge Transfer Partnership
NDUC	Northern Doctors Urgent Care
NEAS	North East Ambulance Service
NEEP	North East Escalation Plan
NHS	National Health Service
NPFS	National Pandemic Flu Service
PCT	Primary Care Trust
PTS	Patient Transport Service
QSR	Quality Standard Report
SCD	Social Care Direct
SHA	Strategic Health Authority
SPA	Single Point of Access
UCC	Urgent Care Centre
UHUK	Urgent Health UK Ltd

Contents

1	Introduction	6
1.1	Knowledge Transfer Partnership	6
1.2	Outline of contents	6
2	Introduction to NDUC	8
2.1	Introduction to Northern Doctors Urgent Care	8
2.1.1	Patient paths through NDUC	8
2.1.2	NHS Pathways	9
2.1.3	Out-of-hours Dental	9
2.1.4	District Nurses	9
2.1.5	Social Care Direct	9
2.2	Problems in Call Centres	9
2.3	Problems with Data Collection	10
2.4	GP Out-of-hours Service	10
2.4.1	Quality Standard Reports	11
2.4.2	Three-server System	12
2.5	Exploratory Data Analysis	12
2.5.1	Dental	12
2.5.2	Resource allocation	16
2.5.3	Moving forward	16
3	Modelling Call Centre Demand	18
3.1	Data Collection	18
3.1.1	Cleaning the dataset	20
3.1.2	Accessing archived data	22
3.2	Possible Modelling Approches	24
3.3	Linear Regression	24
3.3.1	Stepwise Regression	25
3.3.2	Selection Criterion: AIC	25
3.3.3	Selection Criterion: C_p	25
3.4	Establishing the model	25
3.4.1	List of possible Covariates	25
3.4.2	Assessing daily and monthly effects	27
3.4.3	Easter	31
3.4.4	Other Bank Holidays	35
3.4.5	Christmas and New Year	40
3.4.6	Interactions	48
3.5	Validating the new data	55
3.6	Validating model assumptions	58
3.7	Swine flu	60
3.8	Validating the model over time	63

3.8.1	Changes in <i>CPD</i> over time	63
3.8.2	Periodically updating <i>CPD</i>	66
3.9	Conclusions and Recommendations	66
4	Modelling Call Centre Demand	67
4.1	Establishing the Daily Profile	67
4.1.1	Data Collection	67
4.1.2	Locally Weighted Regression	68
4.2	Establishing the Prediction Error	75
4.2.1	Investigating t errors	75
4.3	Investigating Poisson errors	78
4.4	Monitoring the Daily Profile	82
4.4.1	Changes in <i>Daily profile</i> over time	82
4.4.2	Periodically updating <i>Daily Profile</i>	82
4.5	Conclusions and Recommendations	87
5	The Swine flu Pandemic	88
5.1	The Health Protection Agency	89
5.2	Strategic Health Authority	89
5.3	May Bank Holiday 2009	89
5.4	July 2009	89
5.5	October 2009	91
5.5.1	Seven-day Moving Averages	91
5.5.2	North East Strategic Health Authority	97
5.5.3	North East Escalation Plan	97
5.5.4	SHA Situation Reports	98
6	Performance Monitoring	100
6.1	Call handlers	100
6.2	GPs	101
6.2.1	Monitoring Productivity over Shifts	102
6.2.2	The GP Activity Plot	103
6.2.3	The Triage Queue Picture	106
6.2.4	Monitoring Triage Lengths	106
7	Dissemination of Techniques using RExcel	112
7.1	Motivation	112
7.1.1	Installation and startup	112
7.2	Exception Reporting	115
7.3	Transferring <i>CPD</i> model into RExcel	127
7.4	Conclusions and Recommendations	131
8	Further work	132
8.1	Queues	132
8.2	Government vision for the future of the NHS	134
8.2.1	GP consortia	134
8.2.2	Single Point of Access	135
9	Conclusions	136
	Bibliography	137

Chapter 1

Introduction

The research in this thesis took place through a Knowledge Transfer Partnership (KTP) between Northern Doctors Urgent Care (NDUC) and Durham University, which ran between August 2008 and July 2010. This chapter explains the basic ideas of a KTP, as well as defining specific project objectives, before giving an outline of the thesis for the reader.

1.1 Knowledge Transfer Partnership

Knowledge Transfer Partnerships (KTPs), are projects formed between a business, an academic institution, and a recent graduate enabling a company to access skills and expertise to help develop their business [5]. It is part-funded by a government grant, usually through the Technology Strategy Board, and is designed to provide three-way benefits to all parties. Businesses could see increased performance and profitability, academic institutions may develop new teaching material or publish high quality conference papers, and the graduate gets to spend 10% of their time in training and personal development, as well as gaining the valuable experience of tackling a strategic problem.

Each partnership will have specific project objectives and expectations. The tasks set out for this KTP are detailed below, as stated in the KTP Grant Application and Proposal Form:

- To develop and implement robust statistical processes to underpin quality assurance for, and planning of, GP out-of-hours service delivery through
 1. Understanding the business case and sampling NDUC's data; considering data structures and the meaning of variables and their relationships.
 2. Establishing a data framework and Key Performance Indicators (KPIs); preparing data for cleaning and harvesting.
 3. Implementing routine data summaries against the KPIs to meet reporting needs.
 4. Establishing research hypotheses with staff at NDUC, developing new KPIs and collecting new data.
 5. Carrying out statistical investigations on existing data to meet NDUC's requirement for analysis across its operations.
 6. Implementing final methodology for automated analysis and reporting to underpin decisions on resource allocation.

1.2 Outline of contents

NDUC is a not-for-profit community benefit society providing mainly out-of-hours health care to nearly one million patients in the North East of England. The operations of NDUC are

quite complicated, and there are various different patient paths through their system, as well as different methods of data collection, and so the business and operations of NDUC are discussed in chapter 2.

In chapter 3 we discuss the forecasting problem faced by many call centres: how many calls will arrive into the centre on any particular day in the future (inter-day call demand)? The process of data collection and cleaning is also described here. Chapter 4 moves on to discuss the implementation of scatterplot smoothing techniques to forecast 30 minute arrival rates (intra-day call demand).

In July 2009, England was hit by the swine flu pandemic; with wide spread media coverage from the end of April. This meant there was a huge increase in expected demand for out-of-hours services. Chapter 5 discusses the pandemic more thoroughly, and suggests the use of a seven-day moving average to forecast during these conditions.

Central to a smoothly running operation is the performance of call handlers and GPs. These people have direct contact with patients, and in order to achieve optimum staffing we need to ensure they are being productive. This was particularly difficult among the GPs, as there can be large variation in the types of cases they receive. The monitoring and predictive methods implemented to assess this issue are discussed in chapter 6.

The final stage of our project was disseminating these techniques to be used in-house, by non-specialists. Chapter 7 discusses the software RExcel which allows technical R functions to run in the background of Excel spreadsheets.

Chapter 8 highlights suggestions for further work that could not be completed during our project. For example, how will knowing accurate forecasts and server behaviour help you to simulate queues? Not only this, it discusses recent developments in the government, and the vision for a Single Point of Access (SPA) for all non-emergency calls, which could mean a complete change for GP out-of-hours providers and commissioners.

Chapter 9 concludes the thesis and critically evaluates the project. What techniques were implemented? How effective were they in practice? And what can be learnt by other out-of-hours organisations to benefit the community further?

Chapter 2

Introduction to NDUC

This chapter discusses a range of problems faced by call centres trying to forecast demand, and improve operational performance, as well as introducing the call centre used for this investigation.

2.1 Introduction to Northern Doctors Urgent Care

Northern Doctors Urgent Care (NDUC) is a not-for-profit community benefit society which delivers full out-of-hours services to patients in the North East. It is commissioned and funded by local Primary Care Trusts (PCTs) in four areas: Newcastle, North Tyneside, Northumberland, and South Tyneside. The service provides care for approximately one million patients in these areas.

Alongside the GP out-of-hours service NDUC facilitate district nurse provision, provide a pathways service for the North East Ambulance Service (NEAS), an out-of-hours dental service, a telephone answering service and a call handling service for Social Care Direct (SCD), the out-of-hours social care provider for North Tyneside.

2.1.1 Patient paths through NDUC

Patients calling NDUC out-of-hours can expect to receive one of the following responses:

- Telephone advice from a clinician,
- face-to-face contact with a clinician at one of seven Urgent Care Centres (UCCs) based around the North East,
- arrangements for face-to-face contact in a patient's home,
- referral to an appropriate A&E department,
- advice that they require a 999 ambulance, and subsequent transfer to Ambulance Control,
- connection from the ambulance service to the pathways team,
- connection to the dental service,
- connection to the SCD service, or
- connection to the district nursing team.

2.1.2 NHS Pathways

When a patient calls 999, the patient assessment system, known as NHS Pathways, uses evidence-based outcomes to determine the best care. This means that fewer patients will end up in hospital. Where appropriate, a patient can be referred to NDUC from the North East Ambulance Service (NEAS).

The pathways service has its own dedicated telephone line and team including nursing staff, who triage 999 calls transferred from NEAS that require nursing intervention, and dispatchers, who manage the disposition of the various types of pathways call.

Conversely, GPs based at NDUC also support the ambulance service through NEAS Clinician Support. This will be activated in cases where an ambulance has been deployed and arrived at a patient's home, and found the case to be more appropriate to GP urgent care. NEAS then speak to one of our GPs who offer triage support over the telephone.

2.1.3 Out-of-hours Dental

Operating between 18:30 and 23:00 on weekdays, and 09:00 and 22:00 on weekends and bank holidays, NDUC provides an out-of-hours service for patients with dental problems outside normal surgery hours. This service only covers North of Tyne (Northumberland, North Tyneside, and Newcastle PCT areas) differing from the patch covered in the GP service.

Patients using the dental service phone the dedicated dental out-of-hours telephone line. Their call is triaged by a nurse and, if necessary, they will be referred to a Dentist for further consultation.

2.1.4 District Nurses

Northumberland District Nurses assess, prioritise and offer relevant district nursing care for patients living in Northumberland. This service runs at the same time as the GP out-of-hours service, but has a nurse team separate to those working on pathways, dental, and GP out-of-hours.

2.1.5 Social Care Direct

Call handlers at NDUC take calls for Social Care Direct to be actioned by social workers based at NDUC. In the case where the social worker decides there needs to be a face-to-face visit, one of our drivers will take them to the consultation. This service only covers North Tyneside, and runs at the same time as the GP out-of-hours service.

2.2 Problems in Call Centres

There are two main statistical problems investigated during this study:

1. General Forecasting:

- Which services are used more than others?
- How long do different services take?
- How many calls will arrive on any given day?
- When will these calls arrive throughout the shift?

2. Monitoring staff efficiency:

- Agree performance measurements for call handlers, nurses and GPs.
- Establish the distributions of time taken to deal with particular cases.

2.3 Problems with Data Collection

The function of the four main types of out-of-hours staff is detailed below:

Call handlers	Answer calls and assign a priority to each case. These will be picked up by a nurse or GP
GPs	Deal with all out-of-hours triages and consultations. They also support NEAS through telephone triage
Nurses	Deal with all dental out-of-hours triages. They support GP out-of-hours by triaging routine cases
Drivers	Drive GPs to a patient's home

NDUC record around 200,000 calls, with over 100 fields, every year. Each service is extracted and thrown into one all-encompassing data-dump. Cleaning protocols were established to make sure we were including the correct information needed in each investigation. For example, our commissioners are only interested in GP out-of-hours cases for performance reporting; however, when considering GP staffing levels we need to make sure we include NEAS Clinician Support calls as these will take time away from the GPs triaging GP out-of-hours cases. Similarly, nurses help with routine GP out-of-hours cases, but also triage dental out-of-hours calls, and therefore when investigating nurses we extract only these cases.

As well as this, information may be incorrectly entered into the system by users, and a five year data-set will be large, requiring substantial computer power for manipulation and modelling. In particular we faced the following problems in our data-dump, and we implemented sense checks to ensure cases like these were removed:

- Consultation lengths lasting more than a day
- Calls ending before they started
- Triage occurring before the call enters the system
- Calls having a consultation without being triaged

The data collection process changed in October 2005 with the introduction of a new program Adastra. This is explained in more detail in section 3.1.1, and was the only time the process changed.

2.4 GP Out-of-hours Service

The GP out-of-hours service provides GP cover between 18:30 and 08:00 on weekdays, and all day on weekends and bank holidays.

A patient rings the out-of-hours number and reaches a non-clinical call handler who records some details. Using specific urgency criteria, the patient is assessed at this stage as being urgent or routine. At the end of the call the call handler will advise the patient that a clinician will call them back, within a particular time frame. In most cases, one of three things will happen when the GP calls:

1. Patient's symptoms requires advice over the telephone from a GP.
2. The GP decides face-to-face consultation is necessary and arranges an appointment for the patient at one of the UCCs (with or without transport).
3. The GP decides face-to-face consultation is necessary and arranges for a GP to visit the patient at their home.

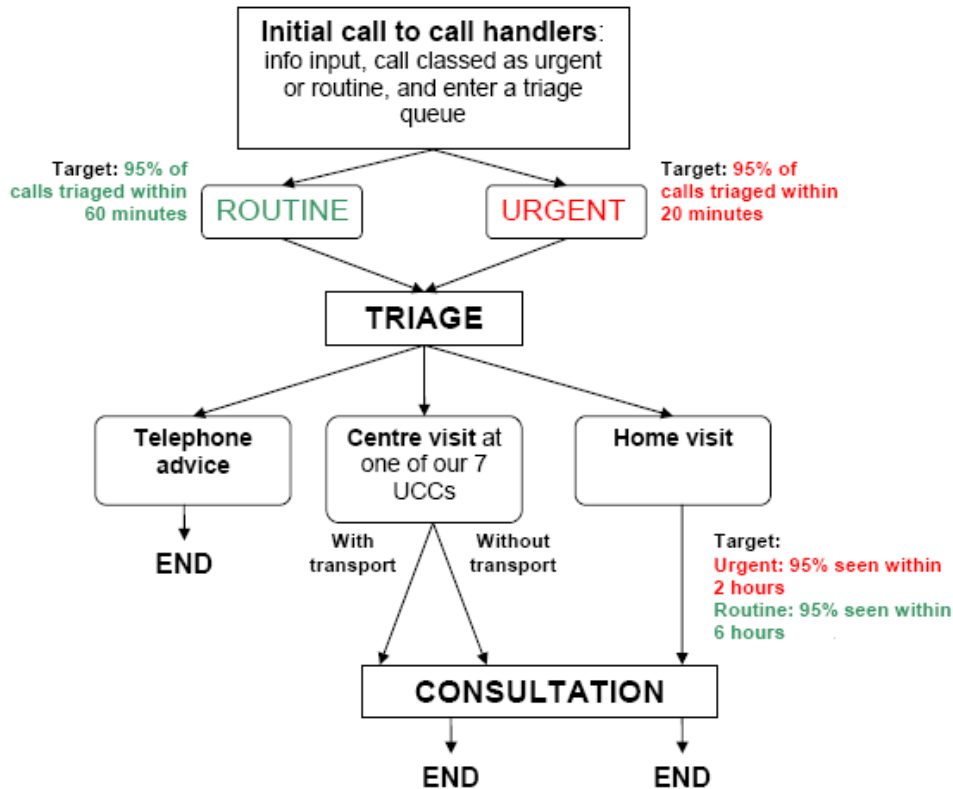


Figure 2.1: Flow chart of basic patient path through the OOH service

This patient path is illustrated in Figure 2.1, along with the KPIs around time frames of waits for patients.

Nearly all cases are triaged by GPs who are based at the call centre, and these triage GPs choose the case outcome of telephone advice, centre visit, or home visit.

For centre visit consultations, there is a GP based at one of the seven UCCs, with 10 minute appointment slots. If the GP has gaps in appointments, they will dip into the triage queue remotely, as and when they can. This may be when the call centre is under pressure and close to missing triage KPIs, while the UCC's KPIs are fine, or when the centres are quiet.

For home visits, one of the GPs based at the call centre will be driven in one of NDUC's fleet vehicles, by a driver, to the patient's home. The length of consultation here is likely to be longer than for centre visits, as these are generally disposed to more complicated cases. At the end of consultation, provided the patient's needs have been met, the GP will either travel back to NDUC base, or to another home visit case.

2.4.1 Quality Standard Reports

NDUC report monthly on its performance to their commissioners with respect to the following Key Performance Indicators (KPIs):

1. Time to triage.
 - Urgent calls triaged within 20 minutes.
 - Routine calls triaged within 60 minutes.
2. Time to Centre Visit assessment

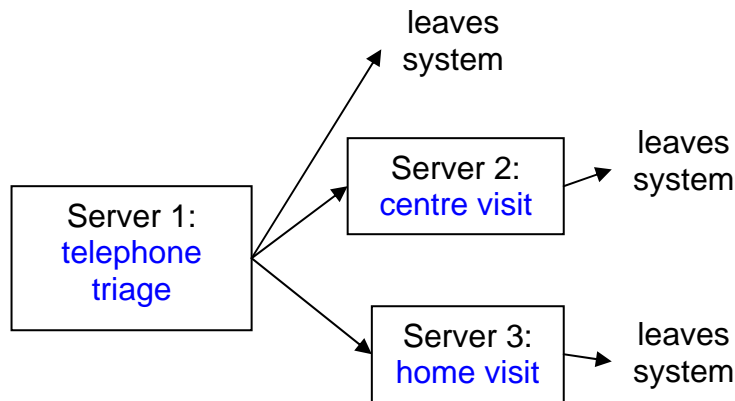


Figure 2.2: The sequential system of NDUC

- Urgent cases seen within 2 hours of the end of telephone triage.
- Routine cases seen within 6 hours of the end of telephone triage.

3. Time to Home Visit assessment

- Urgent cases seen within 2 hours of the end of telephone triage.
- Routine cases seen within 6 hours of the end of telephone triage.

With the above targets NDUC will report as being:

- fully compliant if they achieve 95% or higher,
- partially compliant if they achieve between 90% and 95%, or
- non compliant if they achieve less than 90%.

2.4.2 Three-server System

With respect to the KPIs, NDUC run a three-server system as illustrated in figure 2.2. A patient enters the system waiting for telephone triage (server 1), and then either receives telephone advice (leaves the system), waits for a centre visit appointment (server 2), or waits for a home visit consultation (server 3). To analyse such a system, and keep track of the KPIs in section 2.4.1, we need to keep track of the number of patients at server 1, server 2 and server 3.

2.5 Exploratory Data Analysis

2.5.1 Dental

Before analysing the GP out-of-hours service, we investigated the smaller dental service, to give us an idea of what we faced, and motivate the main work of this thesis. We took a section of data, and without knowing much about the service, tried to summarise its demand.

The dataset consisted of all dental calls arriving into NDUC in the financial year 2007-08. There were 7,160 such calls in this time, and we began by investigating hourly call volume.

Let $\mathbf{h} = (h_0, h_1, \dots, h_{22}, h_{23})$ denote the total calls arriving in each hour (00:00-00:59, 01:00-01:59, ..., 23:00-23:59), and n be the number of distinct days over the whole dataset. In this case, $n = 366$, as 2008 was a leap year and our date range ran from 1st April 2007 until 31st March 2008. Define the daily average, \bar{h} , by:

$$\bar{h} = \frac{\mathbf{h}}{n} \quad (2.1)$$

Next we assess the difference from the overall hourly average for eight different days; Mondays, Tuesdays, ..., Saturdays, Sundays, Bank Holidays. For each day i , ($i = 1, \dots, 8$), define $\mathbf{h}_i = (h_{i,0}, h_{i,1}, \dots, h_{i,22}, h_{i,23})$ as the total calls arriving in each hour on that particular day, and n_i as the number different occurrences of that particular day. Define the daily hourly averages, \bar{h}_i , $i = (1, \dots, 8)$, by:

$$\bar{h}_i = \frac{\mathbf{h}_i}{n_i} \quad (2.2)$$

Apply the same process to assess the difference from the overall hourly average for different months, to understand seasonal variation. For each month, define $\mathbf{h}_j = (h_{j,0}, h_{j,1}, \dots, h_{j,22}, h_{j,23})$, where $j = (1, \dots, 12)$, as the total calls arriving in each hour on that particular month, and n_j as the number days of that particular month. Define the monthly hourly averages \bar{h}_j , $j = (1, \dots, 12)$, by:

$$\bar{h}_j = \frac{\mathbf{h}_j}{n_j} \quad (2.3)$$

We explore how the call arrivals behave with respect to day of week by considering figure 2.3. This figure shows the difference between the daily hourly averages, and the overall hourly average ($\bar{h}_i - \bar{h}$). It seems that Saturday mornings, on average, are busier than Bank Holiday mornings, which are busier, on average, than Sunday mornings. Friday evenings, on average, are busier than any other evening. Referring back to the comment dental out-of-hours (section 2.1.3) the service operates between 18:30 - 23:00 on weekdays, and 09:00 - 22:00 on weekends and bank holidays. There is no contribution of calls coming in between 09:00 - 18:30 on weekdays, and the shape of the graph is not unsurprising.

We apply the same method to see how month affects the service. Figure 2.4 shows the difference between the monthly hourly averages and the overall hourly average ($\bar{h}_j - \bar{h}$), split into four plots containing the months of Spring, Summer, Autumn, and Winter. Summer and Autumn, are generally lower than average, and Winter and Spring are generally above average. Some may be surprised with the ‘busy Spring’ effect, but it should be noted that when you define Spring in this way (months March, April and May), it will always contain three Bank Holidays; Easter and two May Bank Holidays. This is likely to influence why Spring is busier, rather than the season.

Although these are very crude explorations of the data, they were still capable of revealing aspects which had never been seen in NDUC before. These summaries motivated the method used for forecasting GP out-of-hours calls, which is discussed in greater detail in chapter 3.

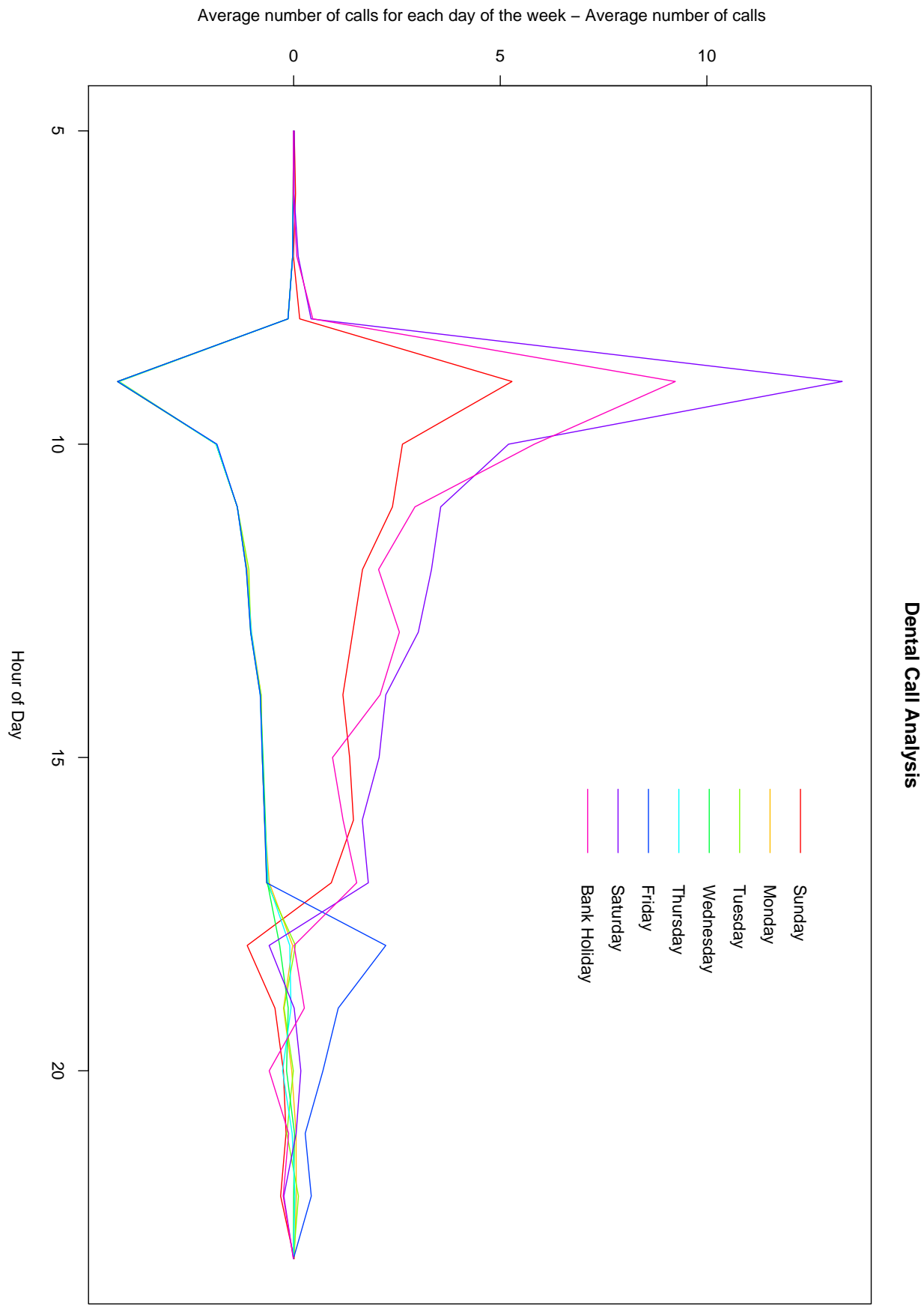


Figure 2.3: Understanding day of week effects

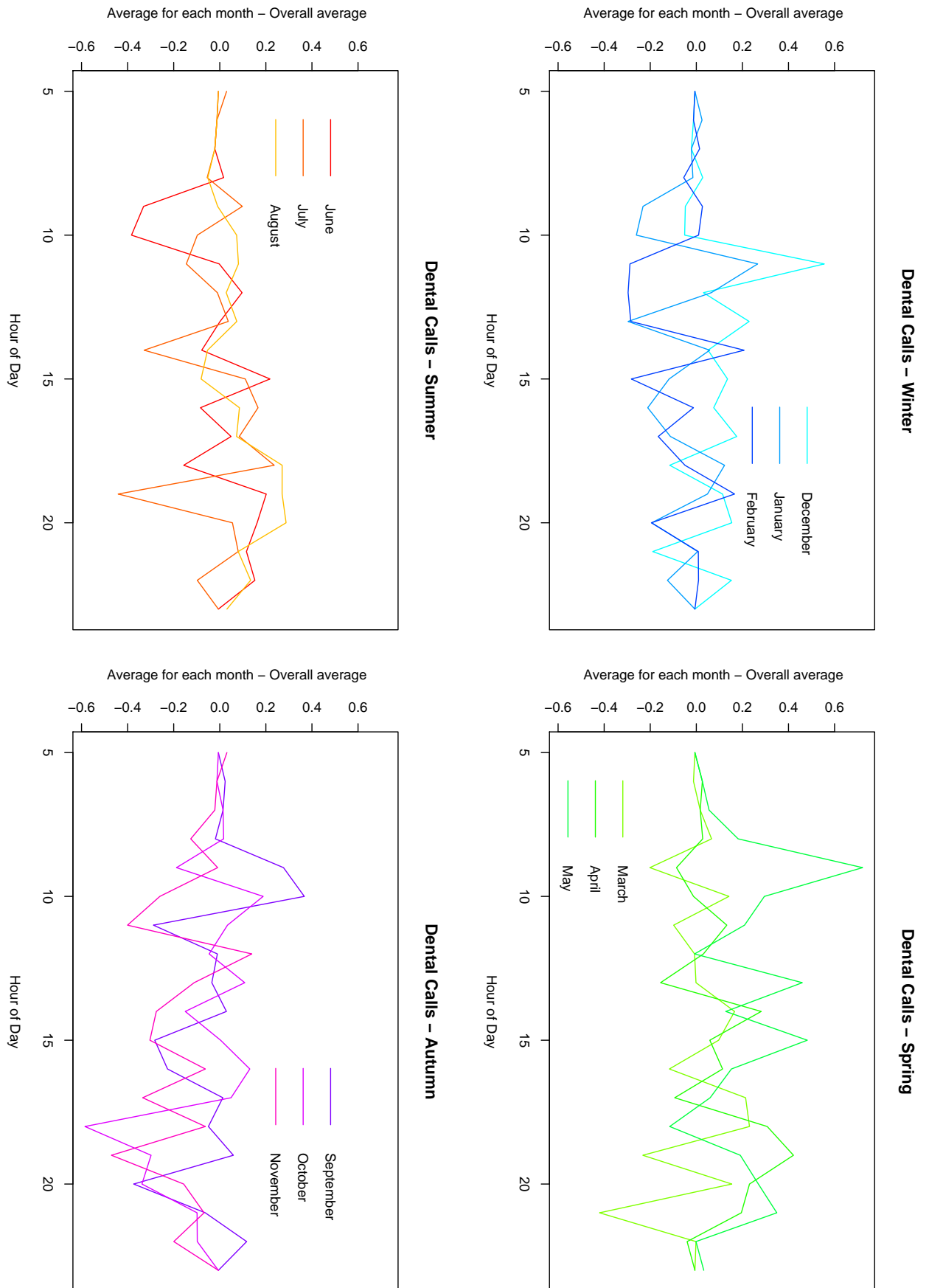


Figure 2.4: Understanding seasonal effects

2.5.2 Resource allocation

How should NDUC staff to meet demand? To motivate ideas about how we should approach such a question, consider the home visit consultation stage, or server 3 in figure 2.2. Here, a patient's call has been triaged by a GP, and that GP has told them that someone will visit their house within a certain time. Then another GP will be driven to the patient's home, anywhere in Northumberland, North Tyneside, Newcastle, or South Tyneside.

This in itself has its own complications. The number of vehicles in our fleet is fixed at 15, so there can never be more than 15 visits occurring at any one time. Geographical factors need to be taken into account: Northumberland is rural, and patients living in Berwick are about 60 miles from our call centre; the River Tyne separates South Tyneside from our headquarters, so travelling there means we must use the tunnel or bridge, which are often congested.

Ignoring the complications for now, let's model this purely in a demand-based way. Consider figure 2.5, which shows a timeline of a Saturday shift running from 08:00 that morning until 07:59 the next day. In what follows, consider the graph between 08:00 - 23:59. The blue line is a model fit of home visit consultations that will arrive in any hour¹. This peaks at about 11 visits between 10:00-12:59, decreasing until 15:59, before slowly levelling off. Let's now assume that one GP can consult with one home visit case in an hour.

Next look at the orange and black solid lines in the figure, which represent the number of home visit GPs and drivers on shift, respectively. You should notice that there are always more drivers on shift than GPs to do home visits. There should always be a spare driver in the case of car breakdowns, or collisions. At 16:00 something strange seemed to happen: the number of visiting GPs decreased and the number of drivers stayed the same, causing an unnecessary gap in the resource. There is a back up of potential GPs that can be pulled from triage to do home visits (dashed orange line), but that doesn't happen often in practice. Therefore understanding the business case and being able to identify a wasteful use of resources would be very useful to NDUC.

2.5.3 Moving forward

We have briefly discussed the issues we are likely to face when undertaking this research, and have also given an argument for the necessity of this work. Firstly, it is essential for out-of-hours healthcare providers to have accurate forecasting, allowing for complicated seasonal and holiday components, that they may be unable to produce through being a small company subject to tight funding. Secondly, they need visual aids that are quick and easy to produce to help them with resource allocation. Most staff at NDUC were unaware of the overstaffing of drivers illustrated in figure 2.5, because the drivers rota was tweaked on several different occasions, resulting in a mismatch.

In what follows we take a deeper look into these issues, and make suggestions for practical and transferrable solutions to GP out-of-hours providers.

¹The model fit is obtain from applying the *Daily Profile* model set up in chapter 4 to arrival rates of home visits at NDUC

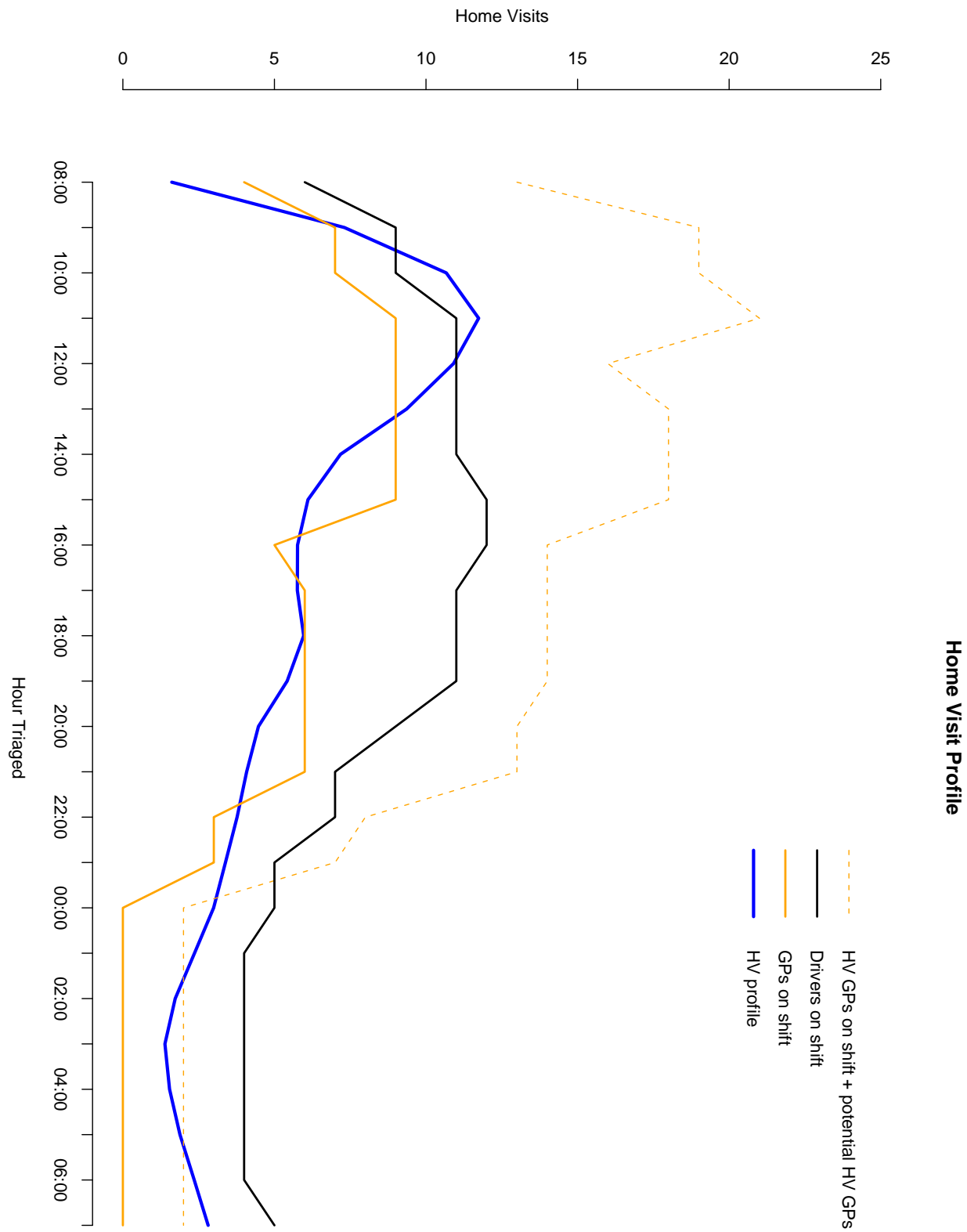


Figure 2.5: Identifying mismatches in resource

Chapter 3

Modelling Call Centre Demand

Understanding call volume at NDUC is essential for the operations team, and this chapter introduces the methods used to address this problem. Before we begin, we split the investigation into two parts:

1. Inter-day (day-to-day) call demand
2. Intra-day (within-day) call demand

This chapter will focus on the first problem, with the latter being explored in chapter 4.

It may be curious to the reader why we model these separately, and the main reason came from the fact that 1. and 2. served different purposes, and were addressed at different stages of the project. Day-to-day call demand would run from midnight to midnight, and would be used externally to illustrate activity levels. Within-day call demand would run over shifts and was designed as an internal communication, to help with staffing levels. It could be argued, however, that inter-day and intra-day call demand are correlated; indeed Avramidis *et al.*[7] investigate the property that there is correlation between arrival counts during parts of the day, as well as dependency between arrival counts on successive days. For this research, it didn't matter if we detached these, provided the models worked well independently. Therefore we will ensure both models adjust appropriately for seasonal and holiday effects, and progress with them separately.

3.1 Data Collection

The data was extracted from the database at NDUC, and is every out-of-hours case since January 1st 2005. This refers to cases in our contracted PCT areas that end up as:

- Telephone advice,
- a centre visit (with or without transport (PTS)),
- a home visit, or
- an A&E Referral (cases referred to our UCCs from A&E departments for GP consultation)

The R output below shows a summary of the dataset (`dprof.2`) running between 1st January 2005 and 1st September 2010, which we will be using to forecast activity into the call centre. You can see there are several fields identifying date, day of week, month, and year as well as dummy variables flagging holiday effects like Easter and Christmas.

```
> summary(dprof.2)
      constype      month      Date      dow
A&E Referral    : 52202  Min.    : 1.000  Min.    : 1.00  Min.    :1.000
```

Casualty Call	: 22253	1st Qu.: 3.000	1st Qu.: 8.00	1st Qu.:1.000
Centre Visit	:185735	Median : 6.000	Median :16.00	Median :4.000
Home Visit	:115811	Mean : 6.252	Mean :15.81	Mean :4.115
PTS Centre Visit:	18731	3rd Qu.: 9.000	3rd Qu.:24.00	3rd Qu.:7.000
Telephone Advice:	200248	Max. :12.000	Max. :31.00	Max. :7.000

hour	hoursplit	doy	year
Min. : 0.00	Min. : 1.000	1 : 3766	2005:102043
1st Qu.: 9.00	1st Qu.: 4.000	360 : 3697	2006:105055
Median :14.00	Median : 7.000	2 : 3233	2007:109745
Mean :13.98	Mean : 6.505	361 : 3174	2008:103536
3rd Qu.:19.00	3rd Qu.: 9.000	84 : 3059	2009:108081
Max. :23.00	Max. :12.000	3 : 3052	2010: 66520

(Other):574999

bankhol	hourq	hourh	realdate
Mode :logical	Min. :1.000	Min. :1.000	27.12.2008: 1156
FALSE:562983	1st Qu.:2.000	1st Qu.:1.000	18.7.2009 : 1088
TRUE :31997	Median :3.000	Median :2.000	3.1.2005 : 1077
NA's :0	Mean :2.502	Mean :1.504	29.12.2007: 1006
	3rd Qu.:3.000	3rd Qu.:2.000	26.3.2005 : 996
	Max. :4.000	Max. :2.000	19.7.2009 : 985
			(Other) :588672

realtime	xmas	nyd	xmasbh
18.2 : 27704	Mode :logical	Mode :logical	Mode :logical
19.1 : 25479	FALSE:592780	FALSE:591214	FALSE:568576
19.2 : 23661	TRUE :2200	TRUE :3766	TRUE :26404
20.1 : 22231	NA's :0	NA's :0	NA's :0
20.2 : 21399			
9.1 : 20723			
(Other):453783			

easterbh	easterfri	eastersat	eastersun
Mode :logical	Mode :logical	Mode :logical	Mode :logical
FALSE:576365	FALSE:590417	FALSE:589615	FALSE:590709
TRUE :18615	TRUE :4563	TRUE :5365	TRUE :4271
NA's :0	NA's :0	NA's :0	NA's :0

eastermon	otherbh	bhsat	bhsun
Mode :logical	Mode :logical	Mode :logical	Mode :logical
FALSE:590564	FALSE:560518	FALSE:582620	FALSE:583997
TRUE :4416	TRUE :34462	TRUE :12360	TRUE :10983
NA's :0	NA's :0	NA's :0	NA's :0

bhmon	maybhsat	maybhsun	maybhmon
Mode :logical	Mode :logical	Mode :logical	Mode :logical
FALSE:583861	FALSE:586459	FALSE:587310	FALSE:587262
TRUE :11119	TRUE :8521	TRUE :7670	TRUE :7718
NA's :0	NA's :0	NA's :0	NA's :0

augbhsat	augbhsun	augbhmon	we
Mode :logical	Mode :logical	Mode :logical	Mode :logical
FALSE:591141	FALSE:591667	FALSE:591579	FALSE:234773
TRUE :3839	TRUE :3313	TRUE :3401	TRUE :360207
NA's :0	NA's :0	NA's :0	NA's :0

```

wd
Mode :logical
FALSE:360207
TRUE :234773
NA's :0

```

We began with these restrictions as these are the cases used in the quality assessment of our service with respect to the six KPIs introduced in section 2.4.1. Every year NDUC records around 100,000 such out-of-hours cases (as can be seen from the R output under the field `year`).

The data was extracted and saved as a .csv file before exporting it into the statistical program R for analysis [20]. The R system implements a dialect of the S language that was developed by AT&T Bell Laboratories by Rick Becker, John Chambers and Allan Wilks. Versions of R are available free of charge for 32-bit versions of Microsoft Windows, Linux and other Unix systems, and the Macintosh. It is available through the Comprehensive R Archive Network (CRAN) at <http://cran.r-project.org/>.

We now move on to discuss how to build a model that can predict how many calls will arrive into the call centre on any given day. First, you have to have an idea of the sorts of things that will influence call volume. For example, there should be a ‘day of week’ effect because the out-of-hours service runs for 14.5 hours on weekdays, and all day Saturdays, Sundays, and bank holidays. You may also suspect that volumes are higher in winter than summer, due to longer vacations, and fewer incidences of influenza.

3.1.1 Cleaning the dataset

We began exploring the dataset by aggregating the cases over individual dates, and plotting them. Figure 3.1 illustrates the first data extract for calls per day for dates between January 1st 2005 and October 31st 2008 (observations y_1, \dots, y_{1400}). This picture has two distinct patterns: one for weekdays (along the bottom), and one for weekends and bank holidays (along the top). The weekday pattern shows call volume ranges between about 100-200 calls per day, with a slight wave, peaking around winter months. The top pattern is clearly oscillating around winter months, and has two peaks every year; the Christmas fortnight and Easter weekend.

Before continuing with any analysis, there were three things which made us question the cleanliness of this data extract: the notable jump, around October 2005, and possible irregularities around June 2005 and June 2006. Discussions with the IT department highlighted the following reasons to explain these irregularities:

1. In October 2005 a new program, Adastra, was introduced into NDUC as a means of recording details. This meant there were more fields to include, and the exclusion criteria used in Microsoft Access to extract the data, had changed.
2. Servers crashed on shift Thursday 30/06/2005. This meant that cases were recorded on paper on this day, and input into the system on Friday 01/07/2005.

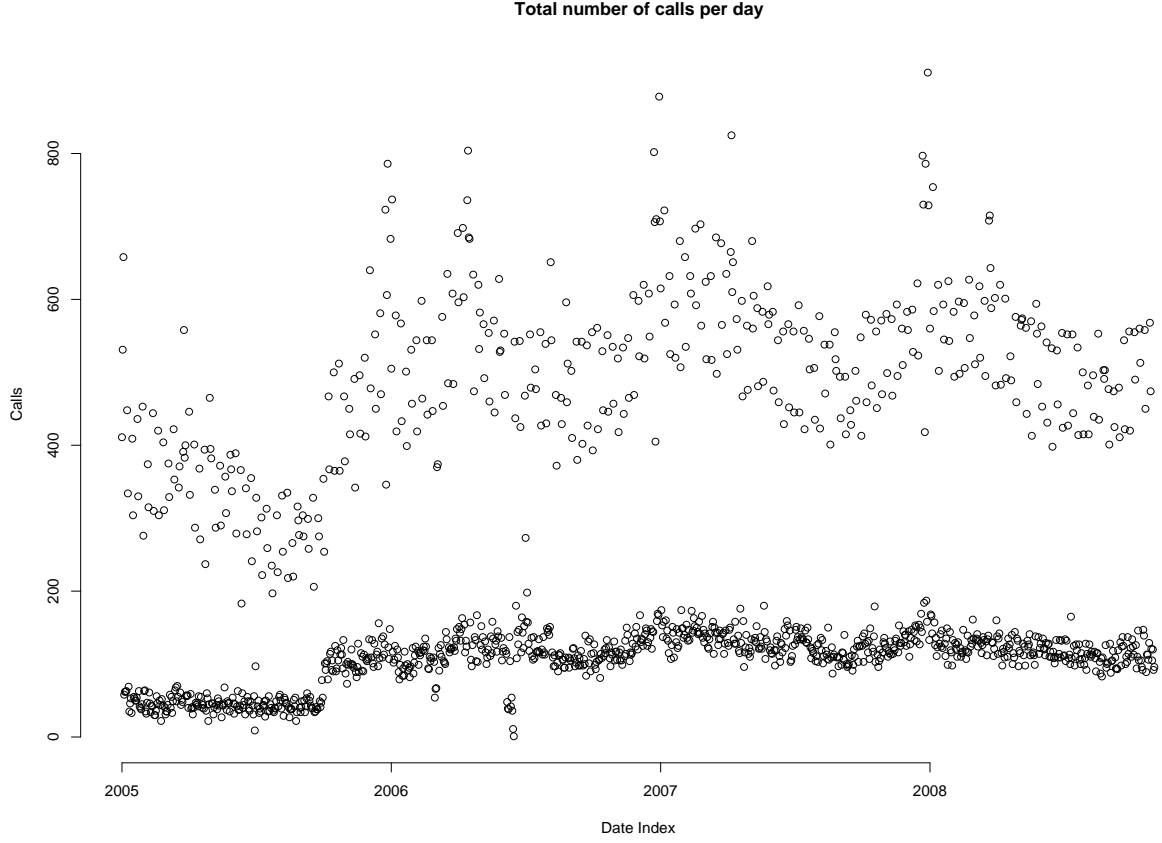


Figure 3.1: Calls per day between 01/01/2005 and 31/10/2008

3. There were server problems in June 2006 for two weeks, resulting in inaccurate calls per day between Sunday 10/06/2006 and Wednesday 05/07/2006.

The following process was put in place to create a second dataset to use for analysis:

1. The data was extracted from the database again, with updated exclusion commands in Microsoft Access.
2. Observations y_{181} (Thursday 30/06/2005) and y_{182} (Friday 01/07/2005) were replaced.
3. Observations y_{526}, \dots, y_{551} (Sunday 10/06/2006 to Wednesday 05/07/2006) were replaced.

The first part of the cleaning process is self explanatory. The second and third parts used averages and linear regression as the method for replacement.

Consider observations y_{181} (Thursday 30/06/2005) and y_{182} , and calculate the aggregate call volume

$$Y = y_{181} + y_{182},$$

over these two days. Y seemed realistic over a normal Thursday and Friday, so we calculated T and F , the sum of calls per day on Thursdays and Fridays, respectively,

$$T = \sum_{\text{Thursdays}} y_i$$

$$F = \sum_{\text{Fridays}} y_i.$$

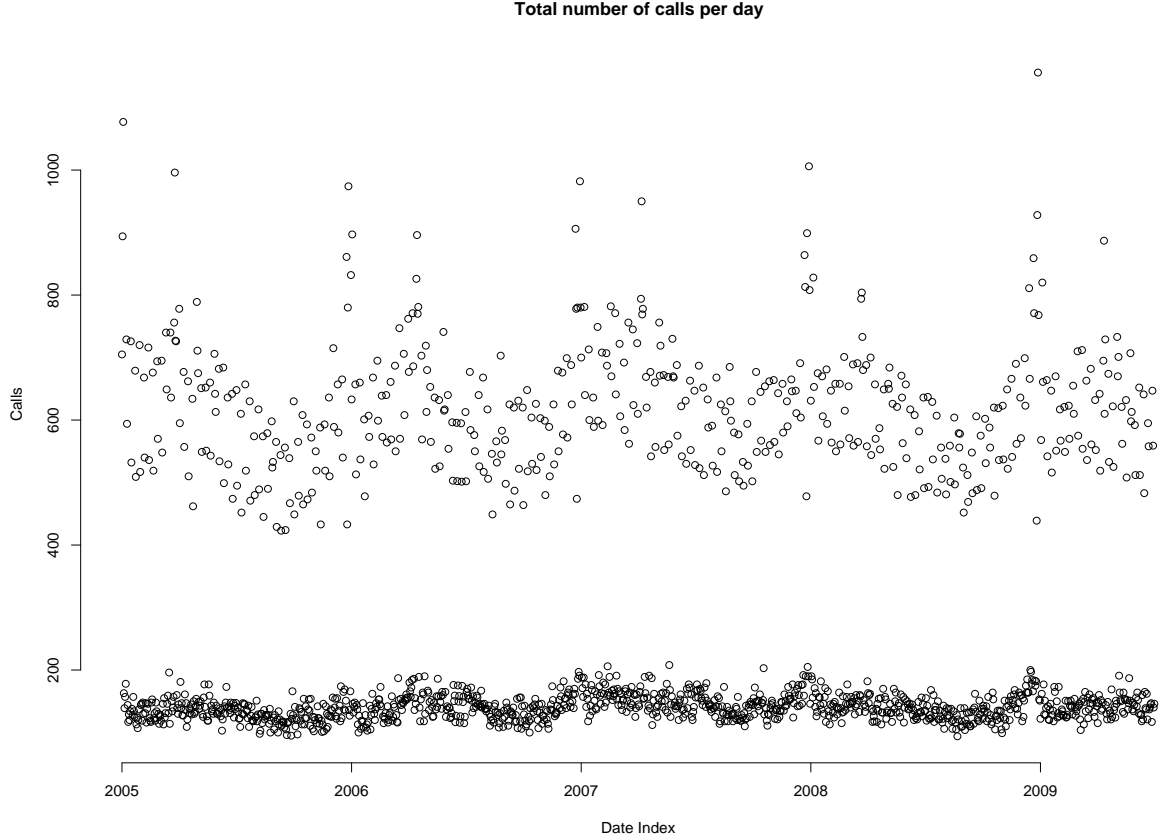


Figure 3.2: Calls per day between 01/01/2005 and 30/06/2009

The proportion $p = T/(T + F)$ was calculated and observations y_{181} , y_{182} were replaced by pY , and $(1 - p)Y$ respectively.

Next, consider observations y_{526}, \dots, y_{551} (Sunday 10/06/2006 to Wednesday 05/07/2006). These 26 observations were summed to give

$$Z = y_{526} + \dots + y_{551},$$

and this seemed realistic over the time period. Given the amount of days in this period, it did not seem appropriate to replace these observations with averages. Therefore we tried back-forecasting, using the model defined in equation 3.21 which is developed later in this chapter. Equation 3.21 was used to produce fits for these days and these were summed to give

$$\hat{Z} = \hat{y}_{526} + \dots + \hat{y}_{551}.$$

The proportion $q = Z/\hat{Z}$ was calculated, and $(y_{526}, \dots, y_{551})$ were replaced with $q(\hat{y}_{526}, \dots, \hat{y}_{551})$.

Figure 3.2 shows the second dataset for calls per day between January 1st 2005 and June 30th 2009 (observations y_1, \dots, y_{1642}). The only calculated replacements of observations remain as described, and the October 2005 jump has disappeared.

3.1.2 Accessing archived data

Before we started forecasting and planning for the Christmas period 2009, we were aware that our model would be restricted by a lack of Christmas holiday observations. We discovered that NDUC archived its monthly Quality Standards Reports (QSRs), which publish monthly

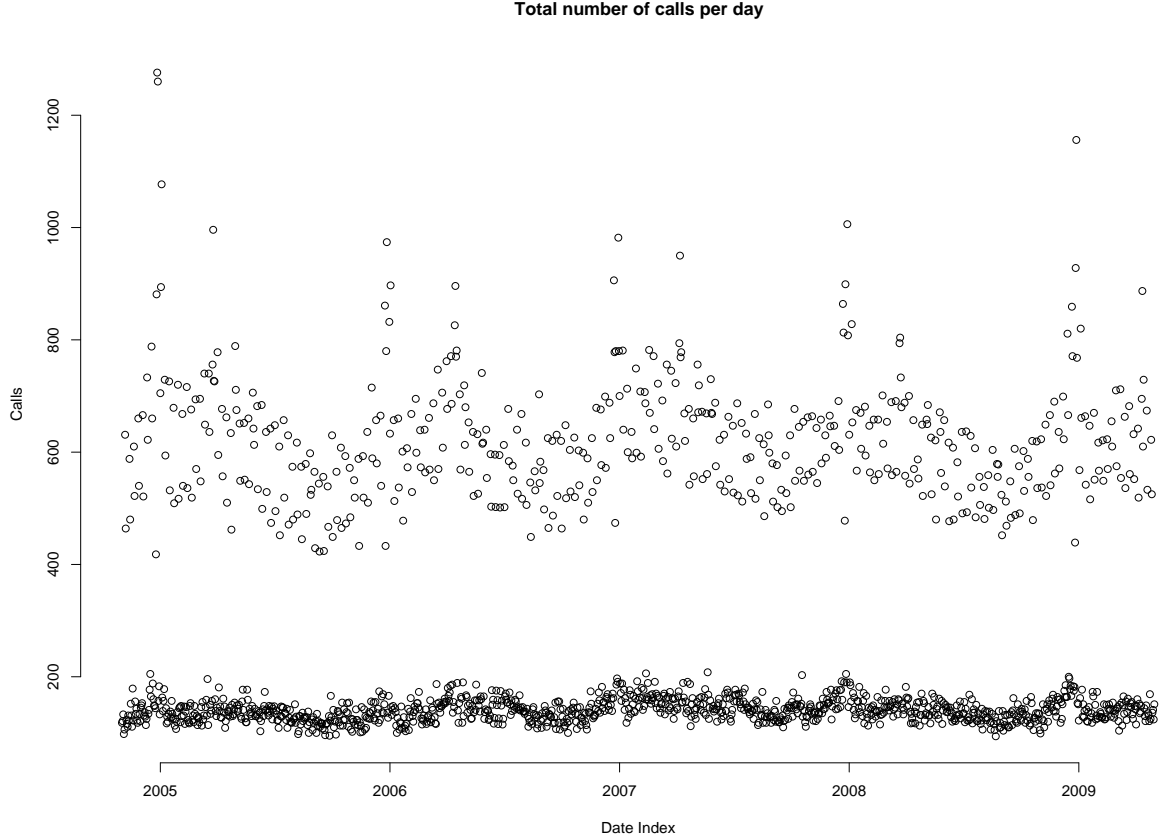


Figure 3.3: Calls per day between 01/11/2004 and 30/06/2009

performance for our commissioners, as well as activity levels, and calls per day. From this we were able to find a table recording calls per day for November 2004 and December 2004 (x_1, \dots, x_{61}). These observations are included in figure 3.3.

Next, we checked the validity of these new observations. To do this we added the new dates to the start of the dataset, observations x_1, \dots, x_{61} , and y_1, \dots, y_{1642} . Then use the latest predictive model to obtain fit errors for $\hat{x}_1, \dots, \hat{x}_{61}$, and $\hat{y}_1, \dots, \hat{y}_{1642}$ (see section 3.4 for the establishment of the model used to generate these). Finally, split the dataset in to two parts:

1. $\mathbf{a} = a_1, \dots, a_{1642}$: the fit errors for $\hat{y}_1, \dots, \hat{y}_{1642}$ (01/01/2005 - 30/06/2009)
2. $\mathbf{b} = b_1, \dots, b_{61}$: the fit errors for $\hat{x}_1, \dots, \hat{x}_{61}$ (01/11/2004 - 31/12/2004)

Let $f_a(\mathbf{a})$ and $f_b(\mathbf{b})$ be the distribution functions of \mathbf{a} and \mathbf{b} , respectively. To check whether the new data is trustworthy consider the null hypothesis (H_0) that the distribution function of \mathbf{b} is equal to the distribution function of \mathbf{a} , that is:

$$H_0 : f_a(\mathbf{a}) = f_b(\mathbf{b})$$

$$H_A : f_a(\mathbf{a}) \neq f_b(\mathbf{b})$$

We test the above hypothesis by using three statistical tests: the two-sample Kolmogorov-Smirnov test, the two-sample t-test, and the two-sample Wilcoxon test (or Mann-Whitney test), to test for the same distribution, means and median respectively. If we fail to reject the null hypothesis, then we would include the new observations in our dataset. The details of this are

discussed in this chapter in section 3.5, after we establish the predictive model in section 3.4. In what follows, the description of building the model assumes we fail to reject the null hypothesis, and use the dataset 01/11/2004 - 30/06/2009.

3.2 Possible Modelling Approches

Alternative approaches to modelling call centre demand are well documented in Shen and Haung's paper [21], which includes discussions on Singular Value Decomposition, ARIMA (Autoregressive Integrated Moving Averages), and Bayesian Gaussian Models. These methods have been applied to core call centre operations, where profitability is often central to the company implementing them, for example, Bank Customer Helplines. The forecasts we generated were not from a regular telephone system, but instead were generated by call handlers. There are clear guidelines about which cases go through as 'real' out-of-hours cases and whether they should be routine or urgent. We are not forecasting calls arriving into our telephone line, but out-of-hours cases being generated by call handlers, so the traditional call centre methods may not be appropriate.

The U.S. Census Bureau's software package for seasonal adjustment, X-12-ARIMA, was also considered, for which documentation can be found online, and in Ladiary and Quenneville's book [17]. On testing the software in November 2008 we could not break it down to forecast to the level that we wanted - daily, and 30 minute arrival rates, so we had to consider another method.

We sought an alternative approach through linear regression. This would be useful as it handles dummy variables easily, and we would have several of these identifying different holiday effects. On the other hand, we could not forget the underlying assumptions of modelling in this way, particularly Normality of errors. Checks would be made throughout the process to ensure this method was appropriate.

3.3 Linear Regression

A linear model assumes a straight line relationship to summarise the dependence of one variable over another. Denote the response variable by Y , which depends on p covariates \mathbf{Z} [15]. In our case the response variable is calls per day, or CPD , and the covariates have to be established. The linear model assumes that for given \mathbf{Z} , a corresponding observation CPD is a combination of $\beta_0 + \beta\mathbf{Z}$ and an error ϵ , that is:

$$CPD = \beta_0 + \beta\mathbf{Z} + \epsilon, \quad (3.1)$$

where β is a vector with p elements. Equation 3.1 is the linear model. $\beta_0 + \beta\mathbf{Z}$ is the model function, and β_0 and β are parameters. Denote the total number of previous observations of CPD as n_{CPD} . We can generate estimates of CPD for $i = 1, \dots, n_{CPD}$ through the equation:

$$\widehat{CPD}_i = \hat{\beta}_0 + \sum_{j=1}^p \hat{\beta}_j z_{ij}, \quad (3.2)$$

from obtaining a least squares estimation of $\beta = \beta_0, \dots, \beta_p$. Estimating the coefficients β in this way means we can estimate the error ϵ by $\hat{\epsilon}_i = CPD_i - \widehat{CPD}_i$. In other words, the residual is equal to the observed value minus the fitted value, for $i = 1, \dots, n$.

From this we can get an unbiased estimate of σ^2 , s^2 , defined below:

$$s^2 = \frac{\hat{\epsilon}_1^2 + \dots + \hat{\epsilon}_n^2}{n - p} = \frac{RSS}{n - p}, \quad (3.3)$$

where RSS stands for the residual sum of squares, n is the number of observations, and p is the number of covariates.

3.3.1 Stepwise Regression

It was clear modelling CPD would depend on a number of covariates, so we used stepwise regression [15] to build the final linear model. Here we choose the Z most correlated with CPD , model $CPD = \beta_0 + \beta_1 Z + \epsilon$, and check if Z is significant. If it's not, adopt the model $CPD = \overline{CPD}$, if it is, search for a second covariate to enter the model. We already know day of week (dow) will have an effect on CPD due to different opening hours, so we begin with a base model:

$$CPD = \beta_0 + \beta_1 dow + \epsilon \quad (3.4)$$

3.3.2 Selection Criterion: AIC

The next problem we faced was deciding what the other relevant covariates were. In practice we need to define a suitable selection criterion that will test whether different terms add value to a model, or take it away.

Akaike defined an information criterion, **AIC** [9], as

$$\mathbf{AIC} = -2 \log(L(\hat{\theta}|y)) + 2K, \quad (3.5)$$

where $L(\hat{\theta}|y)$ is the likelihood of the estimated parameter θ , given observations y , and K is the number of estimable parameters. In the least squares case, using linear regression, this is

$$\mathbf{AIC} = n \log(s^2) + 2K, \quad (3.6)$$

where s^2 is the unbiased estimate of σ^2 , and K is the total number of estimated regression parameters, including the intercept and σ^2 , so $K = p + 2$, where p is the number of covariates.

3.3.3 Selection Criterion: C_p

Consider a submodel CPD_I with p_I covariates, which is different to the full model CPD which has p covariates. Define the residual sum of squares of the submodel as RSS_I . Colin Mallows [18] defined the C_p statistic stated in equation 3.7, where s^2 is the usual estimate of σ^2 in the full model.

$$C_p = \frac{RSS_I}{s^2} + 2p_I - n \quad (3.7)$$

When $p_I = p$, $C_p = p$ [15], so good candidates for submodels are ones where:

1. $C_p \leq p_I$
2. p_I is as small as possible.

The next section shows the steps taken to finalise the model for predicting CPD .

3.4 Establishing the model

3.4.1 List of possible Covariates

Before building the CPD model, table 3.1 states a list of plausible covariates with a short justification to why they should be considered for inclusion to the model. Then we show the thought process behind stepwise selection of the model.

Covariate	Explanation	Justification
<i>dow</i>	Day of week	Out-of-hours service runs for 14.5 hours on weekdays and 24 hours on weekends and we suspect that Mondays are busier than other weekdays.
<i>month</i>	Month	We suspect that summer months will be quieter than winter months due to holidays and low rates of influenza.
<i>year</i>	Year	Are there differences in demand between years?
<i>time</i>	Chronological date	Are there changes in demand over time? For example do patient's attitudes change to the out-of-hours service after publicity campaigns or high profile news stories?
<i>easterbh</i>	Easter weekend	Busier than normal Fridays-Mondays in March and April.
<i>easterfri</i>	Good Fridays	Busier than normal Fridays in March and April.
<i>eastersat</i>	Easter Saturday	Busier than normal Saturdays in March and April.
<i>eastersun</i>	Easter Sunday	Busier than normal Sundays in March and April.
<i>eastermon</i>	Easter Monday	Busier than normal Mondays in March and April.
<i>eastmarch</i>	Easter holidays in March	Expect Easters in March to be busier than April as the weather tends to be colder.
<i>otherbh</i>	Bank Holiday weekends in May and August	We expect these days to be busier than normal Saturdays-Mondays in May and August.
<i>bhsat</i>	Bank Holiday Saturdays	Busier than normal Saturdays in May and August.
<i>bhsun</i>	Bank Holiday Sundays	Busier than normal Sundays in May and August.
<i>bhmon</i>	Bank Holiday Mondays	Busier than normal Mondays in May and August.
<i>maybhsat</i>	Bank Holiday Saturdays in May	Busier than normal Saturdays in May.
<i>maybhsun</i>	Bank Holiday Sundays in May	Busier than normal Sundays in May.
<i>maybhmon</i>	Bank Holiday Mondays in May	Busier than normal Mondays in May.
<i>augbhsat</i>	Bank Holiday Saturdays in August	Busier than normal Saturdays in August.
<i>augbhsun</i>	Bank Holiday Sundays in August	Busier than normal Sundays in August.
<i>augbhmon</i>	Bank Holiday Mondays in August	Busier than normal Mondays in August.
<i>maybh1</i>	First May Bank Holiday weekend	Busier than other Bank Holiday weekends.
<i>xmaswd</i>	Christmas Day falls on a weekday	Busier than a normal weekday in December.
<i>xmaswe</i>	Christmas Day falls on a weekend	Different to a normal weekend in December.
<i>xmasbhwd</i>	Christmas holidays on weekdays	Public holidays on weekdays will be busier than normal weekdays in December and January.
<i>xmasbhwe</i>	Christmas holidays on weekends	Busier than normal weekends in December and January
<i>nydwd</i>	New Years Day falls on a weekday	Busier than a normal weekday in January.
<i>nydwe</i>	New Years Day falls on a weekend	Different to a normal weekend in January.
<i>xmas4day1</i>	Days following Christmas when its a four-day holiday cycle	Unknown influence of the difference between a four-day Christmas holiday versus a two-day Christmas holiday.

Table 3.1: List of plausible covariates to predict CPD

3.4.2 Assessing daily and monthly effects

Equation 3.4 is the base model and we now assess if it is advantageous to include any other terms in the model. To assess the seasonal effect define *month*, to denote the twelve months in a calendar year (January, ..., December) and let *year* denote the years in the dataset (2004, 2005, 2006, 2007, 2008, and 2009). Note that *dow*, *month*, and *year* are factors, denoting dummy variables for each category. Finally, define *time* as the chronological date factor starting from 01/11/2004 until the end of the dataset.

Consider the three linear models defined in equations 3.8, 3.9, and 3.10.

$$CPD = \beta_0 + \beta_1dow + \beta_2month + \epsilon \quad (3.8)$$

$$CPD = \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \epsilon \quad (3.9)$$

$$CPD = \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \epsilon \quad (3.10)$$

In R, you can use the function `step`, to select a linear model by AIC in a stepwise algorithm. To start with, define the linear model of the starting point, in this case equation 3.4. Next define the range of models to be examined in the stepwise search. This can be done with a list containing components `lower`, equation 3.4, and `upper`, equation 3.10. Next you control the argument `scale` which defines the AIC statistic used for selecting the models. In this instance we use the estimate of σ^2 from the largest model, that is s^2 as generated from equation 3.10. Defining the function in this way is equivalent to using the C_p statistic in our selection process. Finally, we specify the `direction` as being “forward”.

It is important to comment here on the choice of our direction. It is well documented [15] that forward selection is less effective than “both” or “backward”, which can be chosen when using `step` in R. This is because forward selection ensures that covariates added to the model, are not subsequently removed, and therefore does not guarantee a good model. Since we had no starting point, and the “forward” selection is relatively simple, we progressed in this way, performing other checks as the model was built.

The R output below shows the output for the function `step` applied as described above. It starts with the model in equation 3.4, and assesses how adding *month*, *year*, *time*, or no term affects the model. In the first step *month* gives the lowest value of $C_p = 30.251$. Therefore we add *month* into the model and move to the next step. Step 2 (where $AIC = 30.25$) now assesses whether adding *year*, *time*, or no term affects the model. *time* and *year* both have a lower value of C_p and we add in *time* and move to the next step. Finally, *year* gives a lower C_p than adding no term, so the base model moves from equation 3.4 to equation 3.10.

```
Start:  AIC=201.42
```

```
CPD ~ dow
```

	Df	Sum of Sq	RSS	Cp
+ month	11	1814593	17766226	30.251
+ year	6	156953	19423866	196.711
<none>			19580819	201.419
+ time	1	12751	19568068	202.061

```
Step:  AIC=30.25
```

```
CPD ~ dow + month
```

	Df	Sum of Sq	RSS	Cp
+ time	1	37726	17728500	28.235
+ year	6	125078	17641148	28.937
<none>			17766226	30.251

Step: AIC=28.24
CPD ~ dow + month + time

	Df	Sum of Sq	RSS	Cp
+ year	6	143120	17585380	25.000
<none>			17728500	28.235

Step: AIC=25
CPD ~ dow + month + time + year

This decision is, of course, supported by the Analysis of Variance table, shown below in the R output. Terms *year*, and *time* are statistically significant at the 5% level, and *month* is extremely influential. Therefore adding these terms to the model will be worthwhile, assuming we trust the p values. Equation 3.10 remains as the base model.

Analysis of Variance Table

```
Model 1: CPD ~ dow
Model 2: CPD ~ dow + month
Model 3: CPD ~ dow + month + year
Model 4: CPD ~ dow + month + year + time
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	1890	19580819				
2	1879	17766226	11	1814593	17.5607	< 2e-16 ***
3	1873	17641148	6	125078	2.2191	0.03877 *
4	1872	17585380	1	55768	5.9367	0.01492 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Below shows the R output summarising the model in equation 3.10. The baseline day of week, month and year are Monday, January and 2004, respectively. Day of week is most significant, with Monday being the busiest weekday, and Wednesday being the quietest compared to it. Saturdays and Sundays are significantly busier than Mondays. January is the busiest month with February, and June - September being significantly quieter at the 1% level of significance, which we would certainly expect in the summer months. Finally, year and time effects are only significant at the 10% level and have conflicting suggestions. Compared to 2004 the volume has been decreasing, year after year. However, there appears to be a very small, and perhaps negligible, increase as time continues to move on. Noting this unusual characteristic, we will ignore it for now, and see if adding further terms to the model addresses this.

Call:
lm(formula = CPD ~ dow + month + year + time, data = cpd)

Residuals:

Min	1Q	Median	3Q	Max
-321.96	-39.03	-10.87	14.79	1032.62

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-6055.5520	3311.8497	-1.828	0.06766	.
dow[T.Tue]	-52.1841	8.7426	-5.969	2.91e-09	***
dow[T.Wed]	-62.8096	8.7520	-7.177	1.07e-12	***
dow[T.Thu]	-56.7820	8.7518	-6.488	1.14e-10	***
dow[T.Fri]	-47.8778	8.7521	-5.470	5.17e-08	***
dow[T.Sat]	461.9110	8.7521	52.777	< 2e-16	***
dow[T.Sun]	354.4505	8.7517	40.501	< 2e-16	***
month[T.Feb]	-36.7717	13.7276	-2.679	0.00746	**
month[T.Mar]	-28.4441	19.2083	-1.481	0.13884	
month[T.Apr]	-33.9393	26.3280	-1.289	0.19754	
month[T.May]	-48.6681	33.8427	-1.438	0.15060	
month[T.Jun]	-115.7716	41.6356	-2.781	0.00549	**
month[T.Jul]	-133.6434	49.6752	-2.690	0.00721	**
month[T.Aug]	-149.2879	57.7313	-2.586	0.00980	**
month[T.Sep]	-187.3714	65.7251	-2.851	0.00441	**
month[T.Oct]	-188.0671	73.7129	-2.551	0.01082	*
month[T.Nov]	-198.0976	81.7105	-2.424	0.01544	*
month[T.Dec]	-115.2010	89.7496	-1.284	0.19947	
year2005	-192.2006	98.3032	-1.955	0.05073	.
year2006	-367.8767	194.9872	-1.887	0.05938	.
year2007	-538.2789	292.0269	-1.843	0.06547	.
year2008	-740.1746	389.3895	-1.901	0.05749	.
year2009	-930.7035	486.6429	-1.912	0.05598	.
time	0.5047	0.2664	1.895	0.05828	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 96.56 on 1679 degrees of freedom

Multiple R-squared: 0.8274, Adjusted R-squared: 0.8251

F-statistic: 350 on 23 and 1679 DF, p-value: < 2.2e-16

Looking at the standardised residual plot in figure 3.4 there is a lack of fit repeating itself year on year around certain holidays. This is not unexpected; during bank holiday weekends our service is open for 24 hours on the Monday, compared to 14.5 hours on a normal weekday. To conclude this section equation 3.10 becomes the base model, and we proceed by searching for holiday effects.

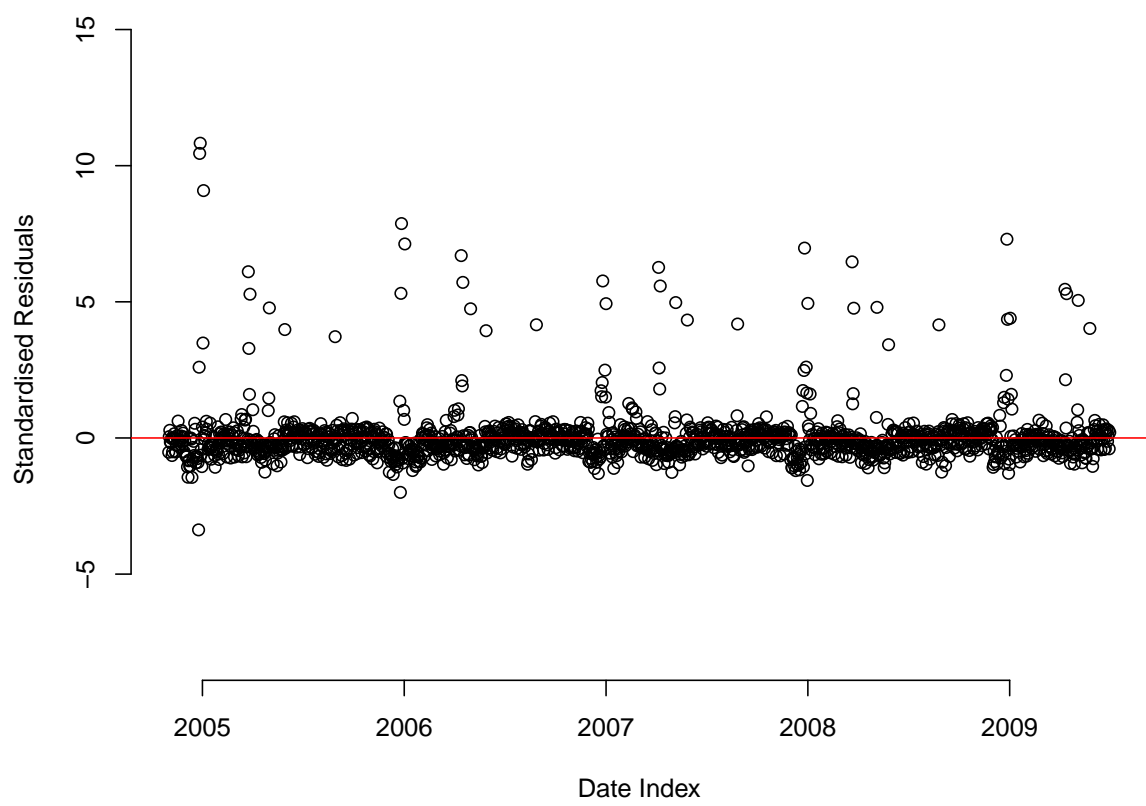


Figure 3.4: Standardised residual plot for equation 3.10 between 01/11/2004 and 30/06/2009

3.4.3 Easter

Easter is a weekend spanning holiday from Good Friday to Easter Monday. Do the Saturdays and Sundays over this holiday behave differently to normal Saturdays and Sundays in March or April? Does it matter whether Easter is in March or April?

Define the following variables to test this:

- *easterbh*: any Friday, Saturday, Sunday or Monday of an Easter weekend between 2005 and 2009.
- *easterfri*, *eastersat*, *eastersun*, and *eastermon*: separate variables for Good Friday, Easter Saturday, Easter Sunday and Easter Monday
- *eastmarch*: True if Easter holidays occur in March.

Note here that it is important to fit the models hierarchically [15], that is, check whether *easterbh* is worth fitting before going down to individual day level *easterfri*, ..., *eastermon*. To do this, define three more linear models:

$$CPD = \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \beta_4easterbh + \epsilon \quad (3.11)$$

$$CPD = \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \epsilon \quad (3.12)$$

$$CPD = \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \beta_9eastmarch + \epsilon \quad (3.13)$$

First assess whether it is worth incorporating a term for the Easter. The R output below shows that including *easterbh* gives a much lower value for C_p (26.0 rather than 392.2), so equation 3.11 then becomes the base model.

Start: AIC=392.2

CPD ~ dow + month + year + time

	Df	Sum of Sq	RSS	Cp
+ easterbh	1	2891604	14693776	26.0
<none>			17585380	392.2

Step: AIC=26

CPD ~ dow + month + year + time + easterbh

Now we need to decide whether it's better to fit to the level of individual days. That is, is equation 3.11 or equation 3.12 a better model? The Analysis of Variance output below comparing equation 3.11 and equation 3.12 shows that equation 3.12 should now become the base model.

Analysis of Variance Table

```
Model 1: CPD ~ dow + month + year + time
+ easterbh
Model 2: CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
  Res.Df    RSS   Df Sum of Sq    F    Pr(>F)
1   1871 14693776
2   1868 13939947    3   753829 33.672 < 2.2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1
```

Finally, we assess whether equation 3.13 is a better model than equation 3.12. That is, does it matter whether Easter is in March or April? The R output from the `step` function gives a lower value of C_p when we do not include *eastmarch* (28.047 rather than 30.000). Therefore equation 3.12 remains as the base model.

```
Start:  AIC=28.05
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
```

	Df	Sum of Sq	RSS	Cp
<none>			13939947	28.047
+ eastmarch	1	353	13939594	30.000

The conclusion for Easter is that we improve the model by fitting individual Easter days, and that Easters in March do not differ considerably to Easters in April, having adjusted for other factors.

Below shows the R output summarising the model in equation 3.12. The day of week, month, time and year effects behave similarly to before, and here you can see that the days of the Easter weekend are significantly busier than normal Fridays, Saturdays, Sundays and Mondays, as expected.

```
Call:
lm(formula = CPD ~ dow + month + year + time + easterfri +
    eastersat + eastersun + eastermon, data = cpd)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-321.247  -33.140   -5.899   15.007 1028.585
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -5392.8267  2903.9437  -1.857  0.06348 .
dow[T.Tue]   -41.0120    7.7051  -5.323 1.16e-07 ***
dow[T.Wed]   -51.4906    7.7137  -6.675 3.35e-11 ***
dow[T.Thu]   -45.5106    7.7134  -5.900 4.38e-09 ***
dow[T.Fri]   -49.6622    7.7528  -6.406 1.94e-10 ***
dow[T.Sat]   468.1730    7.7528  60.387 < 2e-16 ***
dow[T.Sun]   362.3476    7.7525  46.739 < 2e-16 ***
month[T.Feb]  -35.0777   12.0354  -2.915  0.00361 **
month[T.Mar]  -45.7277   16.8614  -2.712  0.00676 **
```

```

month[T.Apr]      -60.8536    23.1198   -2.632   0.00856 **
month[T.May]      -42.1859    29.6739   -1.422   0.15531
month[T.Jun]     -107.5494    36.5071   -2.946   0.00326 **
month[T.Jul]     -123.8422    43.5567   -2.843   0.00452 **
month[T.Aug]     -137.8991    50.6203   -2.724   0.00651 **
month[T.Sep]     -174.1453    57.6298   -3.022   0.00255 **
month[T.Oct]     -173.2904    64.6340   -2.681   0.00741 **
month[T.Nov]     -181.7975    71.6464   -2.537   0.01126 *
month[T.Dec]      -97.0800    78.6955   -1.234   0.21752
year2005         -172.5695    86.1956   -2.002   0.04544 *
year2006         -328.6013   170.9720   -1.922   0.05478 .
year2007         -479.3329   256.0600   -1.872   0.06139 .
year2008         -661.5763   341.4311   -1.938   0.05283 .
year2009         -832.3885   426.7066   -1.951   0.05126 .
time              0.4509      0.2335     1.931   0.05370 .
easterfriTRUE    630.3294    38.5491   16.351   < 2e-16 ***
eastersatTRUE    245.6433    38.5459    6.373  2.39e-10 ***
eastersunTRUE    166.2178    38.5474    4.312  1.71e-05 ***
eastermonTRUE    545.1145    38.5525   14.140   < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

```

```

Residual standard error: 84.65 on 1675 degrees of freedom
Multiple R-squared: 0.8677,    Adjusted R-squared: 0.8656
F-statistic: 406.8 on 27 and 1675 DF,  p-value: < 2.2e-16

```

Figure 3.5 shows the standardised residual plot for equation 3.12, which is an improvement from figure 3.4, but there are clearly several other holiday effects to be taken into account.

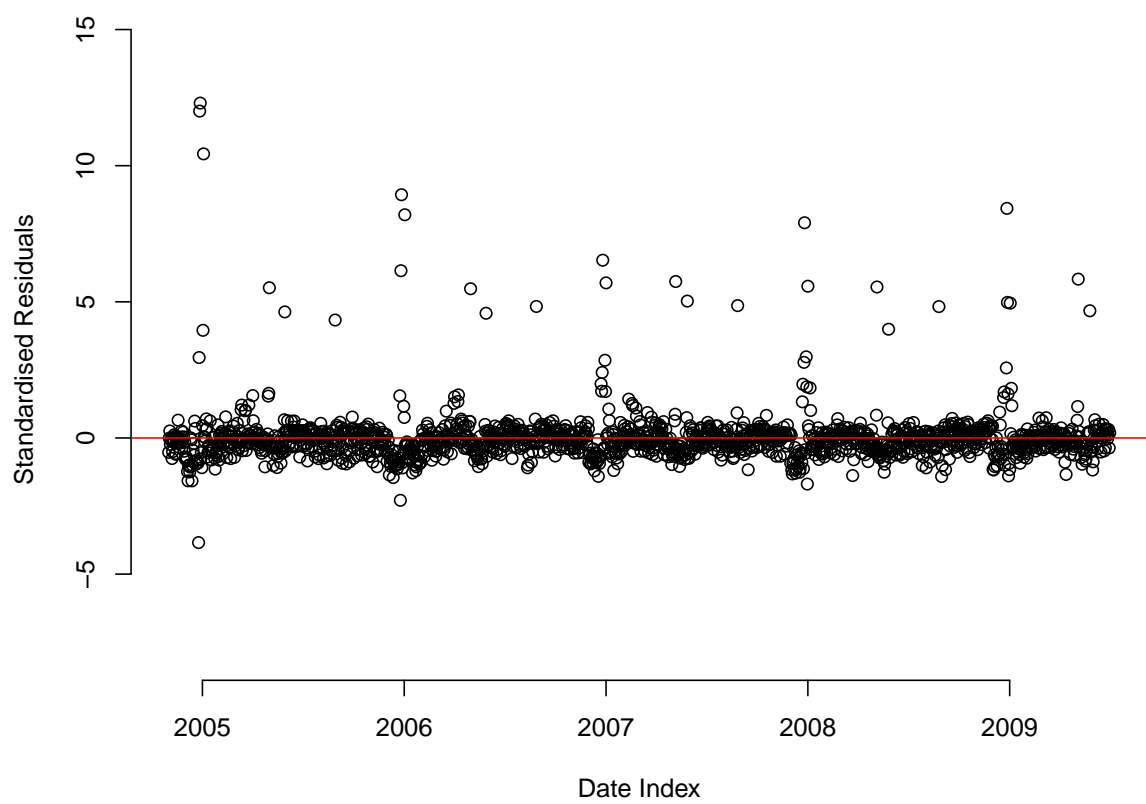


Figure 3.5: Standardised residual plot for equation 3.12 between 01/11/2004 and 30/06/2009

3.4.4 Other Bank Holidays

There are three other bank holiday weekends throughout the year. These three-day weekends span the Saturday, Sunday and Monday of the first Monday in May, the last Monday in May, and the last Monday in August. Define the following variables and consider model equations 3.14 to 3.16:

- *otherbh*: any Saturday, Sunday or Monday of a May or August Bank Holiday weekend between 2005 and 2009.
- *bhsat*, *bhsun*, and *bhmon*: separate variables for Saturdays, Sundays and Mondays, respectively, over May and August Bank Holiday weekends
- *maybhsat*, *maybhsun*, *maybhmon*, *augbhsat*, *augbhsun*, and *augbhmon*: separate variables for Saturdays, Sundays and Mondays, depending whether they are in May or August Bank Holiday weekends
- *maybh1*: The Saturday, Sunday and Monday of the Bank Holiday landing on the first Monday in May.

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9otherbh + \epsilon
 \end{aligned} \tag{3.14}$$

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9bhsat + \beta_{10}bhsun + \beta_{11}bhmon + \epsilon
 \end{aligned} \tag{3.15}$$

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9maybhsat + \beta_{10}maybhsun + \beta_{11}maybhmon + \\
 & + \beta_{12}augbhsat + \beta_{13}augbhsun + \beta_{14}augbhmon + \epsilon
 \end{aligned} \tag{3.16}$$

The R output below shows that *otherbh* is worth fitting, giving a C_p value of 30.00 as opposed to 238.52 when adding no term, and equation 3.14 becomes the base model over equation 3.12.

```

Start:  AIC=238.52
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon

```

	Df	Sum of Sq	RSS	Cp
+ otherbh	1	1412561	12527386	30.00
<none>			13939947	238.52

```

Step:  AIC=30
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ otherbh

```

As with Easter, we needed to assess whether other bank holidays should be fitted to the day level Saturday, Sunday, and Monday, and whether it made a difference if the holidays occurred in May or August. The Analysis of Variance table below shows that equation 3.15 would be the best fit, and we fit to the day level, but not to the month level.

Analysis of Variance Table

```
Model 1: CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ otherbh
Model 2: CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
Model 3: CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ maybhsat + maybhsun + maybhmon +
+ augbhsat + augbhsun + augbhmon
  Res.Df    RSS    Df Sum of Sq      F Pr(>F)
1    1867 12527386
2    1865 10910108     2   1617278 138.4093 <2e-16 ***
3    1862 10878502     3     31606   1.8033 0.1445
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1  1
```

During workforce planning for the August Bank Holiday 2009, I reconsidered the assumption that these three weekends behaved similarly, and checked whether the first May Bank Holiday would have an effect. Define the following linear model to assess this:

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9bhsat + \beta_{10}bhsun + \beta_{11}bhmon + \\
 & + \beta_{12}maybh1 + \epsilon
 \end{aligned}
 \tag{3.17}$$

Equation 3.17 was considered against equation 3.15, and it was better to include *maybh1* than not as illustrated in the R output below (a C_p value of 33 against 37.931). Equation 3.17 now became the base model. The standardised residual plot (figure 3.6) has also improved again.

```
Start:  AIC=37.93
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon

      Df Sum of Sq    RSS    Cp
+ maybh1  1      40420 10869688 33.000
<none>                10910108 37.931
```

```
Step:  AIC=33
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
```

The conclusion for May and August Bank Holidays is that we improve the model by fitting individual weekend days, and the model should also adjust for the first Bank Holiday weekend in May.

The R output below shows the summary for the linear model in equation 3.17, and you can see how it is developing. Now we have included quite a few holiday effects, and the difference of days of week from Monday is not as large, although still exists. A Bank Holiday weekend is busier for every day than a normal weekend, but is only significant for Sundays and Mondays.

Call:

```
lm(formula = CPD ~ dow + month + year + time + easterfri +
    eastersat + eastersun + eastermon + bhsat + bhsun + bhmon +
    maybh1, data = cpd)
```

Residuals:

Min	1Q	Median	3Q	Max
-322.240	-23.172	-3.136	14.280	1028.329

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3672.3817	2570.1718	-1.429	0.153235
dow[T.Tue]	-12.9719	6.8279	-1.900	0.057626 .
dow[T.Wed]	-23.5134	6.8348	-3.440	0.000596 ***
dow[T.Thu]	-17.6143	6.8340	-2.577	0.010038 *
dow[T.Fri]	-21.6936	6.8684	-3.158	0.001615 **
dow[T.Sat]	493.1902	6.9654	70.805	< 2e-16 ***
dow[T.Sun]	385.3655	6.9649	55.330	< 2e-16 ***
month[T.Feb]	-31.1115	10.5438	-2.951	0.003215 **
month[T.Mar]	-37.8972	14.8489	-2.552	0.010793 *
month[T.Apr]	-50.9031	20.4332	-2.491	0.012828 *
month[T.May]	-64.4693	26.2002	-2.461	0.013969 *
month[T.Jun]	-86.4860	32.2767	-2.680	0.007445 **
month[T.Jul]	-98.4602	38.5191	-2.556	0.010672 *
month[T.Aug]	-126.0263	44.5735	-2.827	0.004749 **
month[T.Sep]	-140.2821	50.9817	-2.752	0.005994 **
month[T.Oct]	-134.8497	57.1839	-2.358	0.018480 *
month[T.Nov]	-139.4914	63.3929	-2.200	0.027913 *
month[T.Dec]	-50.4094	69.6346	-0.724	0.469221
year2005	-121.5304	76.2659	-1.594	0.111235
year2006	-226.3892	151.3113	-1.496	0.134795
year2007	-325.8904	226.6244	-1.438	0.150616
year2008	-456.9342	302.1857	-1.512	0.130698
year2009	-576.5041	377.6611	-1.527	0.127072
time	0.3107	0.2067	1.503	0.133003
easterfriTRUE	626.7509	33.6017	18.652	< 2e-16 ***
eastersatTRUE	245.1564	33.6211	7.292	4.70e-13 ***
eastersunTRUE	167.8705	33.6196	4.993	6.56e-07 ***
eastermonTRUE	569.9252	33.6211	16.951	< 2e-16 ***
bhsatTRUE	27.5586	22.5745	1.221	0.222341
bhsunTRUE	61.4703	22.6170	2.718	0.006638 **
bhmonTRUE	454.2798	22.6781	20.032	< 2e-16 ***
maybh1TRUE	60.1034	24.3076	2.473	0.013512 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 73.79 on 1671 degrees of freedom

Multiple R-squared: 0.8997, Adjusted R-squared: 0.8979

F-statistic: 483.6 on 31 and 1671 DF, p-value: < 2.2e-16

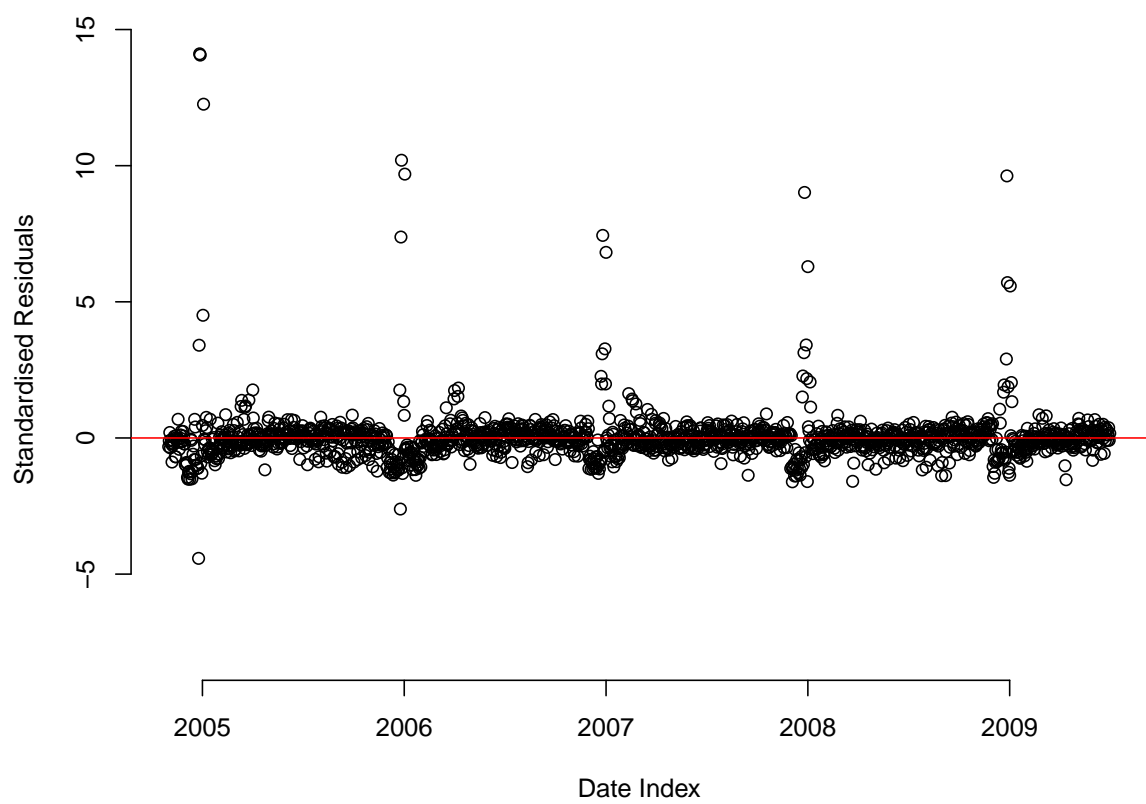


Figure 3.6: Standardised residual plot for equation 3.17 between 01/11/2004 and 30/06/2009

Christmas Day	Four-day cycle?	Cycle	Which day of cycle is Christmas Day?	Historical Information
Monday	✓	Sat-Tue	3	✓2006
Tuesday	×	×	×	✓2007
Wednesday	×	×	×	×
Thursday	✓	Thu-Sun	1	✓2008
Friday	✓	Fri-Mon	1	×
Saturday	✓	Sat-Tue	1	✓2004
Sunday	✓	Sat-Tue	2	✓2005

Table 3.2: The seven cycles of Christmas

3.4.5 Christmas and New Year

There are three public holidays over Christmas: Christmas Day (25th December), Boxing Day (26th December), and New Year’s Day (1st January). Experience within NDUC suggests that the placement of Christmas within the week may have a strong implication for call volumes. For example, when Christmas falls on a Saturday, NDUC operates a ‘four-day’ weekend running from Friday evening until Wednesday morning. If it falls on a Tuesday NDUC only operate from Monday evening until Thursday morning. NDUC expect to see greater call volumes per day if Christmas is on a Saturday compared to a Tuesday. Depending on what day of week Christmas Day falls on, there are seven configurations of this holiday. These are illustrated in table 3.2, which highlights four key points:

1. *Four-day cycle*: Depending on the first day of the configuration, NDUC may be operational for four days continuously. When this happens NDUC are under greater pressure and in-hours administrative tasks, such as preparing the drugs boxes for prescriptions, must take place while the service is live.
2. *Cycle*: Depending when Christmas falls there are three possible four-day cycles; Thursday to Sunday, Friday to Monday, and Saturday to Tuesday. It is suspected that the different cycle may influence call volume.
3. *Which day of cycle is Christmas Day?*: We know that in a four-day cycle Christmas Day will be the quietest. Therefore, depending on when Christmas Day lies in the cycle, this could influence call volume. It is suspected that the busiest cycle is when Christmas Day is the first day of the cycle.
4. *Historical information*: Given the dataset, and the obvious importance of Christmas holidays, there is only one observation of holidays where Christmas lands on a Saturday, Sunday, Monday, Tuesday, and Thursday. There were no observations for Christmas holidays starting on a Wednesday or Friday, and we were trying to fit a Friday configuration in 2009. This emphasises the lack of data for possibly important features.

Define the following variables and use equation 3.18 to assess the Christmas and New Year effect:

- *xmaswe*: True for Christmas Days on Saturdays or Sundays.
- *xmaswd*: True for Christmas Days on weekdays.

- *xmasbhwe*: Saturdays and Sundays falling before and after Christmas Day, which are not Christmas Day or New Years Day
- *xmasbhwd*: Weekdays given as public holidays for Boxing day, or weekdays given in lieu of Christmas Day/Boxing Day falling on weekends.
- *nydwe*: True for New Years Days on Saturdays or Sundays.
- *nydwd*: True for New Years Days on weekdays.

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9bhsat + \beta_{10}bhsun + \beta_{11}bhmon + \\
 & + \beta_{12}maybh1 + \\
 & + \beta_{13}xmasbhwd + \beta_{14}xmasbhwe + \\
 & + \beta_{15}nydwd + \beta_{16}xmaswd + \\
 & + \beta_{17}xmaswe + \beta_{18}nydwe + \epsilon
 \end{aligned} \tag{3.18}$$

The following R output uses the **step** function to show that equation 3.18 should now be the base model instead of equation 3.17.

```

Start:  AIC=5623.38
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1

```

	Df	Sum of Sq	RSS	Cp
+ xmasbhwd	1	6071939	4832065	1466.6
+ nydwd	1	798598	10105406	5078.4
+ xmasbhwe	1	650959	10253045	5179.5
+ xmaswd	1	179307	10724696	5502.6
+ xmaswe	1	131306	10772698	5535.4
+ nydwe	1	5176	10898828	5621.8
<none>			10904004	5623.4

```

Step:  AIC=1466.58
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd

```

	Df	Sum of Sq	RSS	Cp
+ xmasbhwe	1	909198	3922867	845.85
+ nydwd	1	815477	4016587	910.04
+ xmaswd	1	249743	4582322	1297.53
+ xmaswe	1	94980	4737085	1403.53
+ nydwe	1	5781	4826284	1464.62
<none>			4832065	1466.58

Step: AIC=845.85

```
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe
```

	Df	Sum of Sq	RSS	Cp
+ nydwd	1	814232	3108635	290.17
+ xmaswd	1	298976	3623891	643.08
+ xmaswe	1	69311	3853556	800.38
+ nydwe	1	8324	3914543	842.15
<none>			3922867	845.85

Step: AIC=290.17

```
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + nydwd
```

	Df	Sum of Sq	RSS	Cp
+ xmaswd	1	297530	2811105	88.384
+ xmaswe	1	71682	3036953	243.072
+ nydwe	1	11063	3097572	284.591
<none>			3108635	290.168

Step: AIC=88.38

```
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + nydwd + xmaswd
```

	Df	Sum of Sq	RSS	Cp
+ xmaswe	1	67473	2743633	44.170
+ nydwe	1	10599	2800506	83.124
<none>			2811105	88.384

Step: AIC=44.17

```
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + nydwd + xmaswd + xmaswe
```

	Df	Sum of Sq	RSS	Cp
+ nydwe	1	10469	2733164	39.00
<none>			2743633	44.17

```

Step:  AIC=39
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + nydwd + xmaswd + xmaswe + nydwe

```

The R output below shows the summary for the linear model in equation 3.18, and you can see quite a change after the inclusion of Christmas covariates. Monday is still the busiest weekday, but Thursdays are no longer significantly quieter. Month, year and time effects are no longer significant, and only June to November are slightly quieter than January. Holiday effects are highly significant, and quite interesting. For example, if Christmas Day lands on a Saturday or Sunday, this will be quieter than a normal Saturday or Sunday in December, something which may be unexpected to the reader, but is typical in out-of-hours organisations.

Call:

```

lm(formula = CPD ~ dow + month + year + time + easterfri +
    eastersat + eastersun + eastermon + bhsat + bhsun + bhmon +
    maybh1 + xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe,
    data = cpd)

```

Residuals:

	Min	1Q	Median	3Q	Max
	-207.643	-15.658	-0.463	14.927	282.154

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	291.50653	1118.76879	0.261	0.794464
dow[T.Tue]	-9.02705	2.94404	-3.066	0.002203 **
dow[T.Wed]	-9.47566	2.95233	-3.210	0.001355 **
dow[T.Thu]	-3.36873	2.95225	-1.141	0.254004
dow[T.Fri]	-7.39328	2.96693	-2.492	0.012803 *
dow[T.Sat]	500.11378	3.03870	164.582	< 2e-16 ***
dow[T.Sun]	393.98927	3.03122	129.977	< 2e-16 ***
month[T.Feb]	6.30027	4.58120	1.375	0.169241
month[T.Mar]	8.95434	6.44932	1.388	0.165196
month[T.Apr]	5.72717	8.87650	0.645	0.518883
month[T.May]	1.20243	11.38384	0.106	0.915892
month[T.Jun]	-10.05719	14.03076	-0.717	0.473601
month[T.Jul]	-12.13562	16.74668	-0.725	0.468764
month[T.Aug]	-30.75497	19.37981	-1.587	0.112712
month[T.Sep]	-34.27450	22.17083	-1.546	0.122312
month[T.Oct]	-18.89368	24.87039	-0.760	0.447550
month[T.Nov]	-13.41202	27.57397	-0.486	0.626746
month[T.Dec]	20.66169	30.19297	0.684	0.493868
year2005	-2.10957	33.19541	-0.064	0.949336
year2006	10.87844	65.86464	0.165	0.868835
year2007	29.44304	98.64815	0.298	0.765385
year2008	16.16615	131.53898	0.123	0.902201
year2009	13.45832	164.39011	0.082	0.934761
time	-0.01125	0.08998	-0.125	0.900550
easterfriTRUE	621.31399	14.48634	42.890	< 2e-16 ***

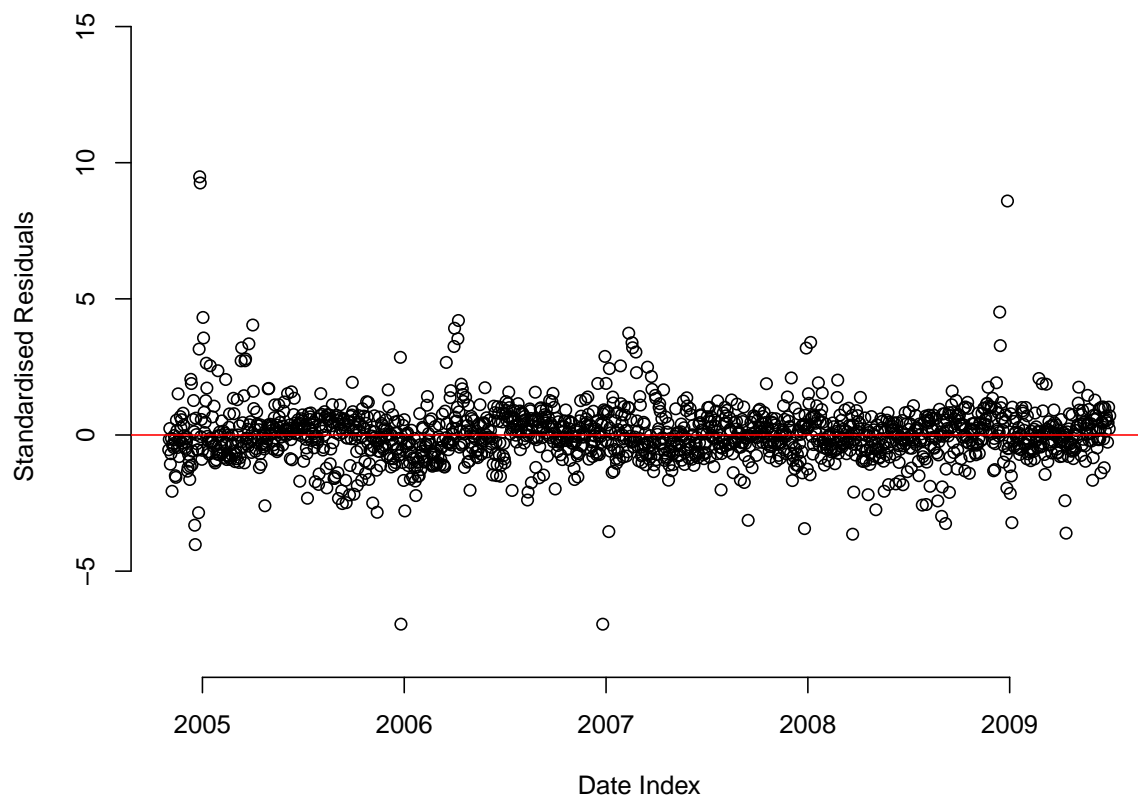


Figure 3.7: Standardised residual plot for equation 3.18 between 01/11/2004 and 30/06/2009

```

eastersatTRUE 247.41818 14.49773 17.066 < 2e-16 ***
eastersunTRUE 168.75393 14.49605 11.641 < 2e-16 ***
eastermonTRUE 579.75445 14.49526 39.996 < 2e-16 ***
bhsatTRUE      34.11890  9.74091  3.503 0.000473 ***
bhsunTRUE      66.00642  9.75668  6.765 1.84e-11 ***
bhmonTRUE     467.11584  9.77893 47.768 < 2e-16 ***
maybh1TRUE     54.32068 10.48148  5.183 2.46e-07 ***
xmasbhwdTRUE  825.38808 10.90146 75.714 < 2e-16 ***
xmasbhweTRUE  221.37770  7.30418 30.308 < 2e-16 ***
xmaswdTRUE    292.80660 18.64611 15.703 < 2e-16 ***
nydwdTRUE     482.02039 18.63521 25.866 < 2e-16 ***
nydweTRUE      71.87613 22.74328  3.160 0.001604 **
xmasweTRUE    -186.92506 22.78288 -8.205 4.58e-16 ***

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 31.81 on 1665 degrees of freedom
Multiple R-squared: 0.9814, Adjusted R-squared: 0.981
F-statistic: 2378 on 37 and 1665 DF, p-value: < 2.2e-16

Although the residual plot (figure 3.7) has considerably improved, the model still fails to predict well over the Christmas period, particularly failing to highlight the importance of the four-day weekend.

Consider the variable *xmas4day1*, which indicates days 2, 3, and 4 of the four day cycle when Christmas Day is the first day of the cycle. This will be true for the three days after Christmas Day when it falls on a Thursday, Friday, or Saturday. Consider the linear model in equation 3.19.

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9bhsat + \beta_{10}bhsun + \beta_{11}bhmon + \\
 & + \beta_{12}maybh1 + \\
 & + \beta_{13}xmasbhwd + \beta_{14}xmasbhwe + \\
 & + \beta_{15}nydwd + \beta_{16}xmaswd + \\
 & + \beta_{17}xmaswe + \beta_{18}nydwe + \\
 & + \beta_{19}xmas4day1 + \epsilon
 \end{aligned} \tag{3.19}$$

The R output below gives a lower value of C_p if we include *xmas4day1* than don't (40.00 against 124.87), and equation 3.19 becomes the base model over equation 3.18. The standardised residual plot for equation 3.19 is illustrated in figure 3.8.

```

Start:  AIC=124.87
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmaswd + nydwd + xmaswe + nydwe + xmasbhwd + xmasbhwe

          Df Sum of Sq    RSS    Cp
+ xmas4day1  1    121267 2611897  40.00
<none>                2733164 124.87

```

```

Step:  AIC=40
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmaswd + nydwd + xmaswe + nydwe + xmasbhwd + xmasbhwe
+ xmas4day1

```

The R output below shows the summary for the linear model in equation 3.19, and you can see the idea of the 'four-day' holiday is clearly important when forecasting for Christmas, and is busier.

```

Call:
lm(formula = CPD ~ dow + month + year + time + easterfri +
    eastersat + eastersun + eastermon + bhsat + bhsun + bhmon +
    maybh1 + xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe +
    xmas4day1, data = cpd)

```

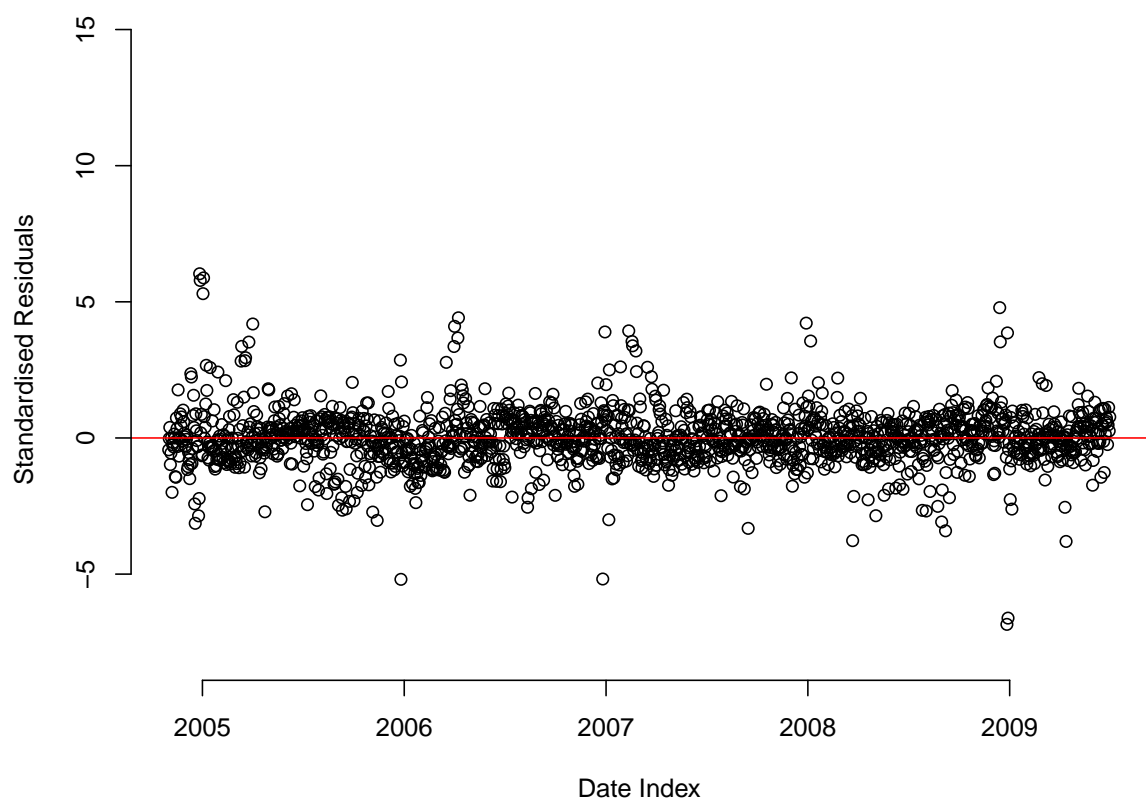



Figure 3.8: Standardised residual plot for equation 3.19 between 01/11/2004 and 30/06/2009

Residuals:

Min	1Q	Median	3Q	Max
-183.1912	-15.4631	-0.4426	15.1281	164.3349

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1032.55476	1065.91738	0.969	0.332835	
dow[T.Tue]	-9.28678	2.80116	-3.315	0.000935	***
dow[T.Wed]	-9.46034	2.80898	-3.368	0.000775	***
dow[T.Thu]	-3.60737	2.80896	-1.284	0.199238	
dow[T.Fri]	-8.15805	2.82346	-2.889	0.003910	**
dow[T.Sat]	500.28720	2.89118	173.039	< 2e-16	***
dow[T.Sun]	393.09858	2.88483	136.264	< 2e-16	***
month[T.Feb]	6.42454	4.35877	1.474	0.140689	
month[T.Mar]	10.83313	6.13781	1.765	0.077750	.
month[T.Apr]	9.43891	8.45016	1.117	0.264151	
month[T.May]	6.72074	10.83911	0.620	0.535313	
month[T.Jun]	-2.68233	13.36111	-0.201	0.840913	
month[T.Jul]	-2.92878	15.94871	-0.184	0.854320	
month[T.Aug]	-19.75506	18.45753	-1.070	0.284640	
month[T.Sep]	-21.38092	21.11679	-1.013	0.311442	
month[T.Oct]	-4.17608	23.68891	-0.176	0.860089	
month[T.Nov]	4.42897	26.26969	0.169	0.866135	
month[T.Dec]	38.38348	28.75811	1.335	0.182157	
year2005	26.76191	31.65879	0.845	0.398052	
year2006	61.54632	62.78332	0.980	0.327082	
year2007	101.89218	94.01766	1.084	0.278630	
year2008	109.00011	125.34836	0.870	0.384657	
year2009	129.42799	156.65321	0.826	0.408805	
time	-0.07121	0.08573	-0.831	0.406288	
easterfriTRUE	621.77391	13.78299	45.112	< 2e-16	***
eastersatTRUE	246.99988	13.79383	17.907	< 2e-16	***
eastersunTRUE	169.45971	13.79229	12.287	< 2e-16	***
eastermonTRUE	579.62950	13.79144	42.028	< 2e-16	***
bhsatTRUE	34.40246	9.26796	3.712	0.000212	***
bhsunTRUE	67.28501	9.28344	7.248	6.45e-13	***
bhmonTRUE	467.43467	9.30415	50.239	< 2e-16	***
maybh1TRUE	53.24841	9.97288	5.339	1.06e-07	***
xmasbhwdTRUE	763.83523	11.36649	67.201	< 2e-16	***
xmasbhweTRUE	196.89552	7.19135	27.380	< 2e-16	***
xmaswdTRUE	294.05780	17.74099	16.575	< 2e-16	***
nydwdTRUE	479.64586	17.73128	27.051	< 2e-16	***
nydweTRUE	68.93834	21.64011	3.186	0.001471	**
xmasweTRUE	-182.67045	21.67904	-8.426	< 2e-16	***
xmas4day1	189.53003	14.31581	13.239	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 30.26 on 1664 degrees of freedom

Multiple R-squared: 0.9832, Adjusted R-squared: 0.9828

F-statistic: 2563 on 38 and 1664 DF, p-value: < 2.2e-16

The conclusion for Christmas Holidays is that we improve the model by fitting individual Christmas, New Years Day, and other holiday days, by weekend and weekday. The model further improves also if we adjust for the four-day cycle starting with Christmas Day.

3.4.6 Interactions

To complete the model, we assessed the validity of including seven interactions, stated in equation 3.20:

1. Does it matter which month weekend days occur? Are weekends in Summer quieter than weekends in Winter? Check this with the interaction between *month* and *we*, where *we* is any Saturday or Sunday.
2. Does it matter which month the days of the week occur? Check this with the interaction between *dow* and *month*.
3. Do months differ as time goes on? Check this with the interaction between *month* and *time*.
4. Do months differ between years? Check this with the interaction between *month* and *year*.
5. Do calls per day for different days of weeks change as time goes on? Check this with the interaction between *dow* and *time*.
6. Do non-Christmas holidays change over time? Check this with the interactions between *time* and *easterbh*, and *time* and *otherbh*.
7. Does it matter which days of the week it is for the three days after the Christmas 4-day cycle? By definition these days will not include Wednesdays or Thursdays, and check this with the interaction between *xmas4day1* and *dow*.

$$\begin{aligned}
CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
& + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
& + \beta_9bh sat + \beta_{10}bh sun + \beta_{11}bh mon + \\
& + \beta_{12}maybh1 + \\
& + \beta_{13}xmasbhwd + \beta_{14}xmasbhwe + \\
& + \beta_{15}nydwd + \beta_{16}xmaswd + \\
& + \beta_{17}xmaswe + \beta_{18}nydwe + \\
& + \beta_{19}xmas4day1 + \\
& + \beta_{20}month : we + \beta_{21}dow : month + \beta_{22}month : year + \\
& + \beta_{23}month : time + \beta_{24}dow : time + \\
& + \beta_{25}time : easterbh + \beta_{26}time : otherbh + \\
& + \beta_{27}dow : xmas4day1 + \epsilon
\end{aligned} \tag{3.20}$$

The R output below suggests we should only include the interactions *month : we*, *dow : xmas4day1*, *month : year*, *month : time*, and *time : easterbh*. This linear model is stated in equation 3.21, with the residual plot illustrated in figure 3.9.

Start: AIC=-2.73

```
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe
+ xmas4day1
```

	Df	Sum of Sq	RSS	Cp
+ month:we	11	269415	1254729	-267.4881
+ dow:month	66	289617	1234527	-178.9904
+ dow:xmas4day1	4	127938	1396207	-130.9015
+ month:year	39	112839	1411305	-44.8310
+ time:easterbh	1	15893	1508251	-17.6429
+ month:time	11	28510	1495634	-11.0727
<none>			1524144	-2.7269
+ time:otherbh	1	361	1523783	-1.1116
+ dow:time	6	9718	1514426	-1.0710

Step: AIC=-267.49

```
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe
+ xmas4day1
+ month:we
```

	Df	Sum of Sq	RSS	Cp
+ dow:xmas4day1	4	128715	1126014	-396.49
+ month:year	39	113396	1141333	-310.18
+ time:easterbh	1	15228	1239501	-281.70
+ month:time	11	29393	1225336	-276.77
<none>			1254729	-267.49
+ time:otherbh	1	148	1254581	-265.65
+ dow:time	6	8232	1246497	-264.25
+ dow:month	55	20202	1234527	-178.99

Step: AIC=-396.49

```
CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe
+ xmas4day1
+ month:we + dow:xmas4day1
```

	Df	Sum of Sq	RSS	Cp
+ month:year	39	115762	1010253	-441.70
+ time:easterbh	1	15286	1110728	-410.76
+ month:time	11	30988	1095026	-407.47

```

<none>                                1126014 -396.49
+ dow:time          6          10051 1115964 -395.19
+ time:otherbh      1           160 1125854 -394.66
+ dow:month         55         19765 1106249 -307.53

```

Step: AIC=-441.7

```

CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe
+ xmas4day1
+ month:we + dow:xmas4day1 + month:year

```

	Df	Sum of Sq	RSS	Cp
+ month:time	11	67647	942606	-491.71
+ time:easterbh	1	7221	1003032	-447.39
<none>			1010253	-441.70
+ dow:time	6	9508	1000744	-439.82
+ time:otherbh	1	7	1010246	-439.71
+ dow:month	55	20191	990061	-353.20

Step: AIC=-491.71

```

CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe
+ xmas4day1
+ month:we + dow:xmas4day1 + month:year + month:time

```

	Df	Sum of Sq	RSS	Cp
+ time:easterbh	1	7577.4	935028	-497.77
<none>			942606	-491.71
+ dow:time	6	9470.7	933135	-489.79
+ time:otherbh	1	75.2	942530	-489.79
+ dow:month	55	20504.2	922101	-403.53

Step: AIC=-497.77

```

CPD ~ dow + month + year + time
+ easterfri + eastersat + eastersun + eastermon
+ bhsat + bhsun + bhmon
+ maybh1
+ xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe
+ xmas4day1
+ month:we + dow:xmas4day1 + month:year + month:time + time:easterbh

```

	Df	Sum of Sq	RSS	Cp
<none>			935028	-497.77
+ time:otherbh	1	84.1	934944	-495.86
+ dow:time	6	9107.7	925920	-495.47

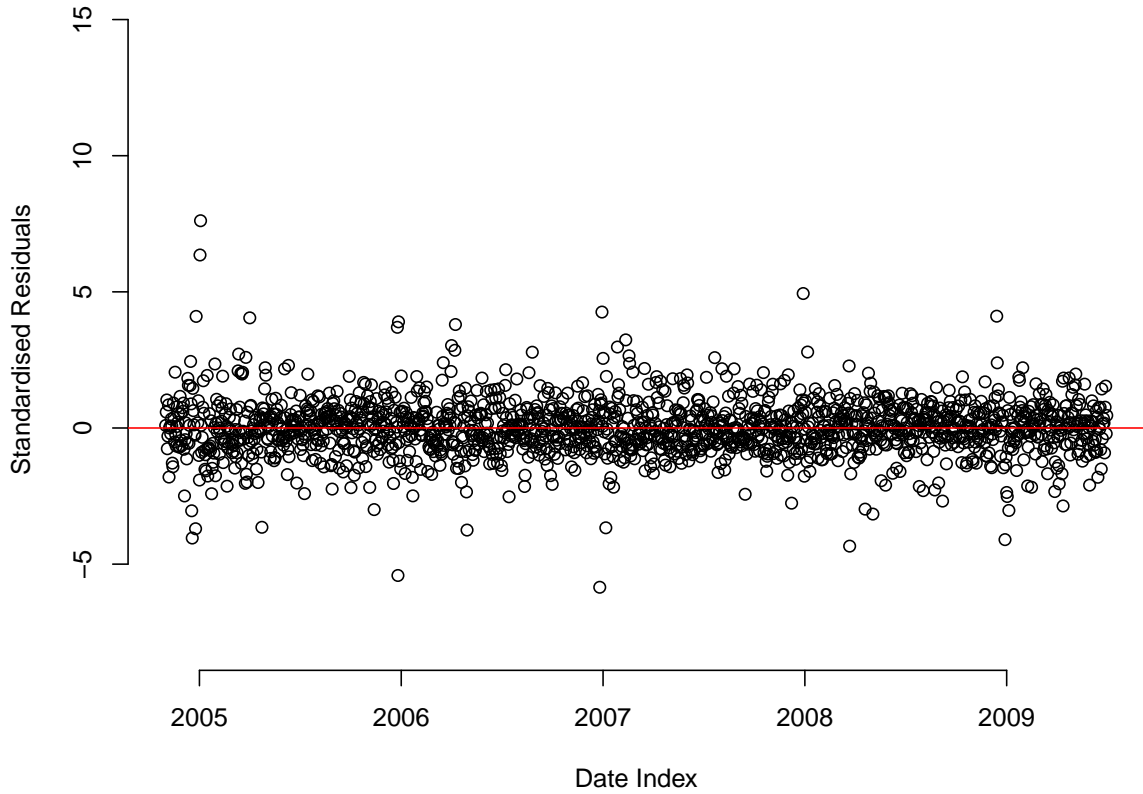


Figure 3.9: Standardised residual plot for equation 3.21 between 01/11/2004 and 30/06/2009

```
+ dow:month      55    20708.3 914320 -409.81
```

The residual plot in figure 3.9 is slightly better, but still inadequate around Christmas.

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9bhsat + \beta_{10}bh sun + \beta_{11}bhmon + \\
 & + \beta_{12}maybh1 + \\
 & + \beta_{13}xmasbhwd + \beta_{14}xmasbhwe + \\
 & + \beta_{15}nydwd + \beta_{16}xmaswd + \\
 & + \beta_{17}xmaswe + \beta_{18}nydwe + \\
 & + \beta_{19}xmas4day1 + \\
 & + \beta_{20}month : we + \beta_{21}dow : xmas4day1 + \\
 & + \beta_{22}month : year + \beta_{23}month : time + \\
 & + \beta_{24}time : easterbh + \epsilon
 \end{aligned} \tag{3.21}$$

The R output below shows the summary for the linear model in equation 3.21. Consider the interaction *dow : xmas4day1*. The possible days of *xmas4day1* are Friday, Saturday, Sunday, Monday and Tuesday. This shows that the Mondays are the busiest, closely followed

by the Tuesdays, with the Fridays, Saturdays, and Sundays being quieter. This captures the expectations of NDUC as they described the worst day of any four-day weekend being the Monday and Tuesday. This is when Christmas Day falls on a Friday or Saturday, and it is believed that many patients expect their surgeries to be open on these days, causing high demand at NDUC when their call is diverted to us. There are also several NA coefficients, illustrating effects that are not estimable, for example, weekends in December.

Call:

```
lm(formula = calls ~ dow + month + year + time + easterfri +
    eastersat + eastersun + eastermon + bhsat + bhsun + bhmon +
    maybh1 + xmasbhwd + xmasbhwe + xmaswd + nydwd + nydwe + xmaswe +
    xmas4day1 + we:month + dow:xmas4day1 + month:year + time:easterbh +
    time:month, data = cpd)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-126.32199	-13.31980	-0.06893	12.81517	163.55573

Coefficients: (19 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1.688e+04	3.299e+03	5.117	3.49e-07	***
dow[T.Tue]	-9.578e+00	2.247e+00	-4.263	2.13e-05	***
dow[T.Wed]	-9.108e+00	2.252e+00	-4.045	5.49e-05	***
dow[T.Thu]	-2.961e+00	2.253e+00	-1.314	0.188958	
dow[T.Fri]	-6.908e+00	2.268e+00	-3.046	0.002360	**
dow[T.Sat]	5.448e+02	5.539e+00	98.361	< 2e-16	***
dow[T.Sun]	4.389e+02	5.531e+00	79.359	< 2e-16	***
month[T.Feb]	-2.253e+04	4.889e+03	-4.609	4.37e-06	***
month[T.Mar]	-2.147e+04	4.594e+03	-4.675	3.19e-06	***
month[T.Apr]	3.872e+03	4.711e+03	0.822	0.411317	
month[T.May]	-1.393e+04	4.847e+03	-2.874	0.004101	**
month[T.Jun]	-1.322e+04	4.682e+03	-2.824	0.004799	**
month[T.Jul]	-5.875e+03	4.715e+03	-1.246	0.212934	
month[T.Aug]	-5.019e+03	4.821e+03	-1.041	0.298063	
month[T.Sep]	-2.396e+04	4.854e+03	-4.935	8.84e-07	***
month[T.Oct]	-1.500e+04	4.732e+03	-3.170	0.001552	**
month[T.Nov]	-2.991e+04	4.397e+03	-6.803	1.44e-11	***
month[T.Dec]	-2.813e+04	5.028e+03	-5.594	2.60e-08	***
year2005	-2.501e+03	4.647e+02	-5.381	8.50e-08	***
year2006	-2.125e+03	4.175e+02	-5.091	3.98e-07	***
year2007	-1.677e+03	3.835e+02	-4.374	1.30e-05	***
year2008	-1.281e+03	3.665e+02	-3.494	0.000488	***
year2009	-8.958e+02	3.691e+02	-2.427	0.015349	*
time	-1.111e+00	2.337e-01	-4.755	2.17e-06	***
easterfriTRUE	1.172e+03	1.540e+02	7.608	4.72e-14	***
eastersatTRUE	7.710e+02	1.540e+02	5.005	6.20e-07	***
eastersunTRUE	6.928e+02	1.541e+02	4.497	7.40e-06	***
eastermonTRUE	1.133e+03	1.541e+02	7.351	3.13e-13	***
bhsatTRUE	5.441e+01	8.090e+00	6.725	2.43e-11	***
bhsunTRUE	8.602e+01	8.170e+00	10.528	< 2e-16	***
bhmonTRUE	4.706e+02	7.735e+00	60.837	< 2e-16	***
maybh1TRUE	4.098e+01	9.129e+00	4.489	7.68e-06	***

xmasbhwTRUE	7.421e+02	1.042e+01	71.203	< 2e-16	***
xmasbhweTRUE	1.614e+02	7.220e+00	22.352	< 2e-16	***
xmaswdTRUE	2.892e+02	1.450e+01	19.949	< 2e-16	***
nydwdTRUE	4.647e+02	1.472e+01	31.579	< 2e-16	***
nydweTRUE	4.099e+01	1.810e+01	2.264	0.023682	*
xmasweTRUE	-2.229e+02	1.822e+01	-12.230	< 2e-16	***
xmas4day1	3.741e+02	2.668e+01	14.020	< 2e-16	***
monthJan:weTRUE	-2.232e+01	6.547e+00	-3.409	0.000668	***
monthFeb:weTRUE	-9.575e+00	6.902e+00	-1.387	0.165592	
monthMar:weTRUE	-3.371e+00	6.822e+00	-0.494	0.621282	
monthApr:weTRUE	-2.544e+01	6.922e+00	-3.676	0.000245	***
monthMay:weTRUE	-5.470e+01	7.137e+00	-7.664	3.11e-14	***
monthJun:weTRUE	-6.099e+01	6.826e+00	-8.935	< 2e-16	***
monthJul:weTRUE	-6.888e+01	7.060e+00	-9.757	< 2e-16	***
monthAug:weTRUE	-8.468e+01	7.300e+00	-11.599	< 2e-16	***
monthSep:weTRUE	-9.496e+01	7.135e+00	-13.308	< 2e-16	***
monthOct:weTRUE	-6.272e+01	7.113e+00	-8.818	< 2e-16	***
monthNov:weTRUE	-5.469e+01	6.829e+00	-8.008	2.23e-15	***
monthDec:weTRUE	NA	NA	NA	NA	
dow[T.Tue]:xmas4day1	-7.315e+00	3.428e+01	-0.213	0.831070	
dow[T.Wed]:xmas4day1	NA	NA	NA	NA	
dow[T.Thu]:xmas4day1	NA	NA	NA	NA	
dow[T.Fri]:xmas4day1	-3.631e+02	3.489e+01	-10.409	< 2e-16	***
dow[T.Sat]:xmas4day1	-1.071e+02	3.671e+01	-2.917	0.003582	**
dow[T.Sun]:xmas4day1	-3.212e+02	3.202e+01	-10.033	< 2e-16	***
month[T.Feb]:year2005	2.287e+03	5.004e+02	4.570	5.26e-06	***
month[T.Mar]:year2005	2.190e+03	4.699e+02	4.660	3.43e-06	***
month[T.Apr]:year2005	-3.958e+02	4.814e+02	-0.822	0.411113	
month[T.May]:year2005	1.414e+03	4.946e+02	2.859	0.004303	**
month[T.Jun]:year2005	1.338e+03	4.773e+02	2.803	0.005127	**
month[T.Jul]:year2005	4.709e+02	3.693e+02	1.275	0.202543	
month[T.Aug]:year2005	4.090e+02	3.771e+02	1.085	0.278289	
month[T.Sep]:year2005	1.864e+03	3.792e+02	4.915	9.79e-07	***
month[T.Oct]:year2005	1.172e+03	3.694e+02	3.173	0.001539	**
month[T.Nov]:year2005	2.115e+03	4.722e+02	4.478	8.05e-06	***
month[T.Dec]:year2005	2.172e+03	3.913e+02	5.551	3.32e-08	***
month[T.Feb]:year2006	1.743e+03	3.755e+02	4.642	3.73e-06	***
month[T.Mar]:year2006	1.668e+03	3.525e+02	4.733	2.41e-06	***
month[T.Apr]:year2006	-2.409e+02	3.612e+02	-0.667	0.504977	
month[T.May]:year2006	1.089e+03	3.712e+02	2.934	0.003398	**
month[T.Jun]:year2006	1.026e+03	3.581e+02	2.865	0.004228	**
month[T.Jul]:year2006	3.685e+02	2.464e+02	1.496	0.134927	
month[T.Aug]:year2006	3.148e+02	2.516e+02	1.251	0.211072	
month[T.Sep]:year2006	1.295e+03	2.530e+02	5.121	3.41e-07	***
month[T.Oct]:year2006	8.205e+02	2.464e+02	3.330	0.000888	***
month[T.Nov]:year2006	1.380e+03	4.495e+02	3.070	0.002173	**
month[T.Dec]:year2006	1.489e+03	2.611e+02	5.703	1.40e-08	***
month[T.Feb]:year2007	1.168e+03	2.505e+02	4.662	3.40e-06	***
month[T.Mar]:year2007	1.091e+03	2.352e+02	4.640	3.78e-06	***
month[T.Apr]:year2007	-1.963e+02	2.410e+02	-0.814	0.415483	
month[T.May]:year2007	6.983e+02	2.476e+02	2.821	0.004850	**

month[T.Jun]:year2007	6.578e+02	2.390e+02	2.753	0.005978	**
month[T.Jul]:year2007	1.716e+02	1.235e+02	1.389	0.164990	
month[T.Aug]:year2007	1.525e+02	1.261e+02	1.209	0.226875	
month[T.Sep]:year2007	6.265e+02	1.268e+02	4.940	8.61e-07	***
month[T.Oct]:year2007	4.068e+02	1.235e+02	3.294	0.001010	**
month[T.Nov]:year2007	5.700e+02	4.578e+02	1.245	0.213241	
month[T.Dec]:year2007	7.242e+02	1.308e+02	5.537	3.58e-08	***
month[T.Feb]:year2008	5.619e+02	1.256e+02	4.474	8.21e-06	***
month[T.Mar]:year2008	5.346e+02	1.179e+02	4.534	6.22e-06	***
month[T.Apr]:year2008	-1.129e+02	1.208e+02	-0.935	0.350133	
month[T.May]:year2008	3.258e+02	1.240e+02	2.627	0.008695	**
month[T.Jun]:year2008	3.082e+02	1.197e+02	2.574	0.010149	*
month[T.Jul]:year2008	NA	NA	NA	NA	
month[T.Aug]:year2008	NA	NA	NA	NA	
month[T.Sep]:year2008	NA	NA	NA	NA	
month[T.Oct]:year2008	NA	NA	NA	NA	
month[T.Nov]:year2008	-2.066e+02	4.955e+02	-0.417	0.676732	
month[T.Dec]:year2008	NA	NA	NA	NA	
month[T.Feb]:year2009	NA	NA	NA	NA	
month[T.Mar]:year2009	NA	NA	NA	NA	
month[T.Apr]:year2009	NA	NA	NA	NA	
month[T.May]:year2009	NA	NA	NA	NA	
month[T.Jun]:year2009	NA	NA	NA	NA	
month[T.Jul]:year2009	NA	NA	NA	NA	
month[T.Aug]:year2009	NA	NA	NA	NA	
month[T.Sep]:year2009	NA	NA	NA	NA	
month[T.Oct]:year2009	NA	NA	NA	NA	
month[T.Nov]:year2009	NA	NA	NA	NA	
month[T.Dec]:year2009	NA	NA	NA	NA	
time:easterbhTRUE	-4.062e-02	1.129e-02	-3.599	0.000330	***
month[T.Feb]:time	1.579e+00	3.424e-01	4.611	4.32e-06	***
month[T.Mar]:time	1.504e+00	3.215e-01	4.679	3.12e-06	***
month[T.Apr]:time	-2.628e-01	3.293e-01	-0.798	0.424932	
month[T.May]:time	9.792e-01	3.384e-01	2.893	0.003863	**
month[T.Jun]:time	9.298e-01	3.266e-01	2.847	0.004475	**
month[T.Jul]:time	4.301e-01	3.371e-01	1.276	0.202098	
month[T.Aug]:time	3.698e-01	3.442e-01	1.074	0.282792	
month[T.Sep]:time	1.712e+00	3.461e-01	4.946	8.37e-07	***
month[T.Oct]:time	1.079e+00	3.371e-01	3.200	0.001403	**
month[T.Nov]:time	2.145e+00	3.267e-01	6.566	6.97e-11	***
month[T.Dec]:time	2.004e+00	3.573e-01	5.609	2.40e-08	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 24.19 on 1598 degrees of freedom

Multiple R-squared: 0.9897, Adjusted R-squared: 0.989

F-statistic: 1475 on 104 and 1598 DF, p-value: < 2.2e-16

3.5 Validating the new data

Given equation 3.21 is the working model for our predictions, use this to confirm the validity of the data between 01/11/2004 - 31/12/2004. Below is the summary of the dataset to predict calls per day, where day of week, month, year and the dummy variables relating to whether the date was a holiday are omitted.

```
> summary(newdata)
      CPD           time
Min.   : 94.0   Min.   :01.11.04
1st Qu.: 132.0   1st Qu.:31.12.05
Median : 151.0   Median :02.03.07
Mean    : 288.5   Mean    :02.03.07
3rd Qu.: 528.5   3rd Qu.:30.04.08
Max.    :1276.0   Max.    :30.06.09
```

After applying equation 3.21 to `newdata`, we get the predictive model `basemodel` in R. The residual plot of `basemodel` is shown in figure 3.10, with residuals for dates 01/11/2004 - 31/12/2004 coloured red, and 01/01/2005 - 30/06/2009 coloured blue, and on first glance, it looks like they could be from the same distribution. The residuals range from a minimum of -126.3 to a maximum of 163.6, and the R output summary of them is detailed below.

```
> summary(basemodel$resid)
      Min.      1st Qu.      Median      Mean      3rd Qu.      Max.
-1.263e+02 -1.332e+01 -6.893e-02  4.275e-16  1.282e+01  1.636e+02
```

Let $\hat{\epsilon}$ be the full set of residuals from `basemodel`, and define **a** and **b** as follows:

1. **a** = a_1, \dots, a_{1642} : $\hat{\epsilon}_{62}, \dots, \hat{\epsilon}_{1703}$ (01/01/2005 - 30/06/2009)
2. **b** = b_1, \dots, b_{61} : $\hat{\epsilon}_1, \dots, \hat{\epsilon}_{61}$ (01/11/2004 - 31/12/2004)

Let m be the size of the sample **a**, in this case 1642, and n be the size of the sample **b**, in this case 61.

The Kolmogorov-Smirnov tests if **a** and **b** are drawn from the same continuous distribution [14]. Consider the problem in which a random sample of m observations, $\mathbf{A} = A_1, \dots, A_m$ is taken from a distribution for which the density function $f_A(\mathbf{a})$ is unknown, and an independent sample $\mathbf{B} = B_1, \dots, B_n$ is taken from another distribution for which the density function $f_B(\mathbf{b})$ is also unknown. Assume that both $f_A(\mathbf{a})$ and $f_B(\mathbf{b})$ are continuous functions, and the following hypothesis is to be tested:

$$H_0 : f_A(x) = f_B(x) \quad \text{for } -\infty < x < \infty, \quad (3.22)$$

$$H_0 : f_A(x) \neq f_B(x) \quad \text{for } -\infty < x < \infty. \quad (3.23)$$

Let $f_a(x)$ denote the sample density function calculated from the observed values A_1, \dots, A_m , and let $f_b(x)$ denote the sample density function calculated from the observed values B_1, \dots, B_n . Next we consider the statistic D_{mn} , the maximum difference between the two sample density functions:

$$D_{mn} = \sup_{-\infty < x < \infty} |f_a(x) - f_b(x)|. \quad (3.24)$$

If H_0 is true and $f_A(x) = f_B(x)$ then the sample density functions $f_a(x)$ and $f_b(x)$ will tend to be very alike. Furthermore, Kolmogorov and Smirnov established the following result in the 1930s:

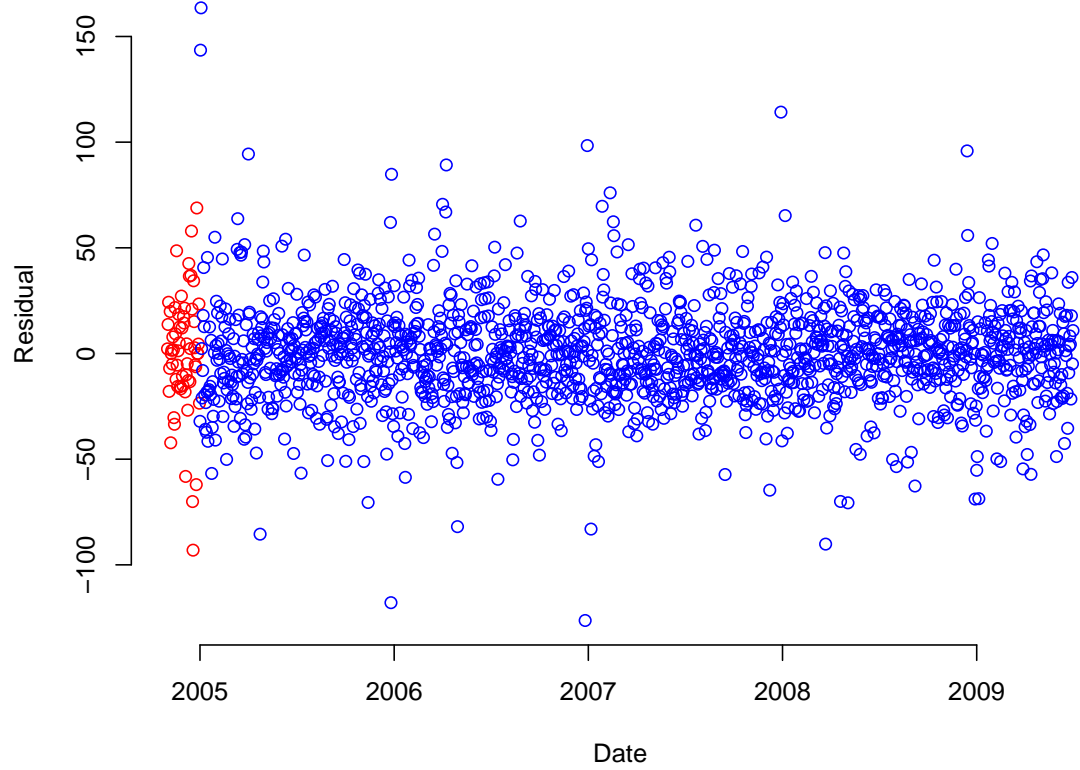


Figure 3.10: Residual plot for equation 3.21 between 01/11/2004 and 30/06/2009

$$\lim_{m \rightarrow \infty, n \rightarrow \infty} \Pr \left[\left(\frac{mn}{m+n} \right)^{1/2} D_{mn} \leq t \right] = H(t), \quad (3.25)$$

where

$$H(t) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} \exp^{-2i^2 t^2}. \quad (3.26)$$

The test procedure specifies that we reject H_0 if

$$\left(\frac{mn}{m+n} \right)^{1/2} D_{mn} \geq c, \quad (3.27)$$

Where $c = H^{-1}(1 - \alpha)$, m and n are large, and α is the level of significance. The values of $H(t)$ can be found in DeGroot and Schervish's book in Table 9.30 [14, pp 570]. If n and m are large and the test is to be carried out at significance level 0.05, we choose $c = H^{-1}(0.95) = 1.36$.

In R, we use the function `ks.test` and run a two-sample, two-sided test for this. In our case the product of n and m is 100,162 which means exact p-values cannot be generated (we require $n \times m < 10,000$ for this), and asymptotic distributions are used. $D = 0.0857$ is the value of the test statistic D_{mn} , as seen in the R output below.

In our case, m and n are large, and we will carry out the test at the 5% level of significance, where $c = 1.36$. Therefore we check equation 3.27 and find that

$$\left(\frac{mn}{m+n} \right)^{1/2} D_{mn} = 0.6570973 < 1.36,$$

and we fail to reject H_0 at the 5% level of significance.

Two-sample Kolmogorov-Smirnov test

```
data: a and b
D = 0.0857, p-value = 0.781
alternative hypothesis: two-sided
```

For completeness we also checked if **a** and **b** had the same means and medians by running the functions `t.test` and `wilcox.test` in R.

Let \bar{a} be the mean of the sample **a**, and \bar{b} be the mean of the sample **b**. We test the hypothesis

$$H_0 : \bar{a} = \bar{b}$$

$$H_A : \bar{a} \neq \bar{b}$$

Below shows the R output from the two-sample t-test, and we fail to reject H_0

Welch Two Sample t-test

```
data: a and b
t = 0, df = 62.914, p-value = 1
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-7.479429 7.479429
sample estimates:
mean of x mean of y
3.862475e-16 1.527723e-15
```

Below shows the R output to test the null hypothesis H_0 : the median of **a** = the median of **b**. With a p-value of 0.6139 we fail to reject H_0

```
Wilcoxon rank sum test with continuity correction

data:  a and b
W = 48178, p-value = 0.6139
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -7.300309  4.348803
sample estimates:
difference in location
 -1.479519
```

We therefore decided it was beneficial to add November and December 2004's data to the dataset, and have done so in forecasting since July 2009.

3.6 Validating model assumptions

By this point, we have fixed the linear model to be equation 3.21, and it is appropriate to make standards checks of our assumptions, namely, Normality of the residuals. Figure 3.11 shows the Normal Quantile plot for the residuals generated from equation 3.21. For the residuals to be Normal, we would see a straight line here. This plot shows that the residual distribution has fatter tails than the normal distribution. As such, we might want to treat any prediction intervals that we develop as conservative, especially in predicting call volumes for days where we feel the model is still not capturing important effects, typically Christmas.

It is also worth making a comment here about sample size. When there is a large sample size, small p values are inevitable. We want to fit models which are good, and justified by circumstances, but we don't want to fit a model which over-tunes to effects which seem inexplicable, even with hindsight. All of the covariates in our model were sensible choices before we started building it (recall table 3.1), and it made sense to staff at NDUC.

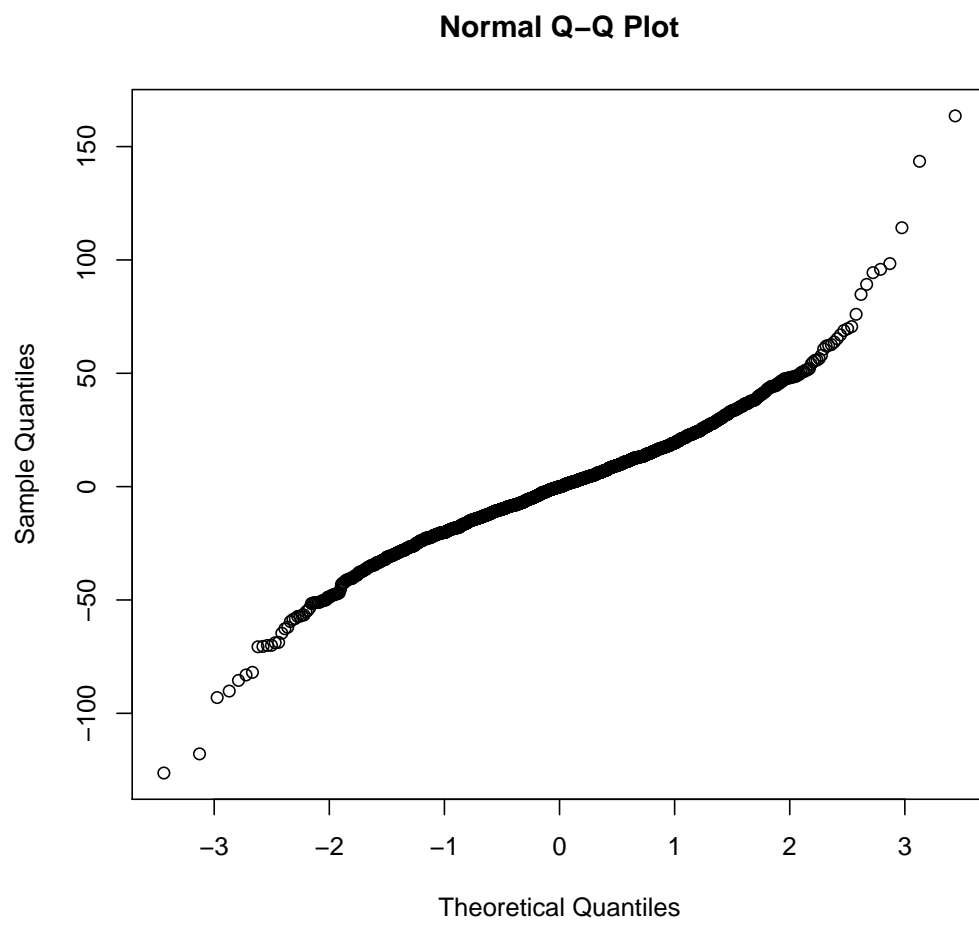


Figure 3.11: Normal quantile plot of residuals from equation 3.21

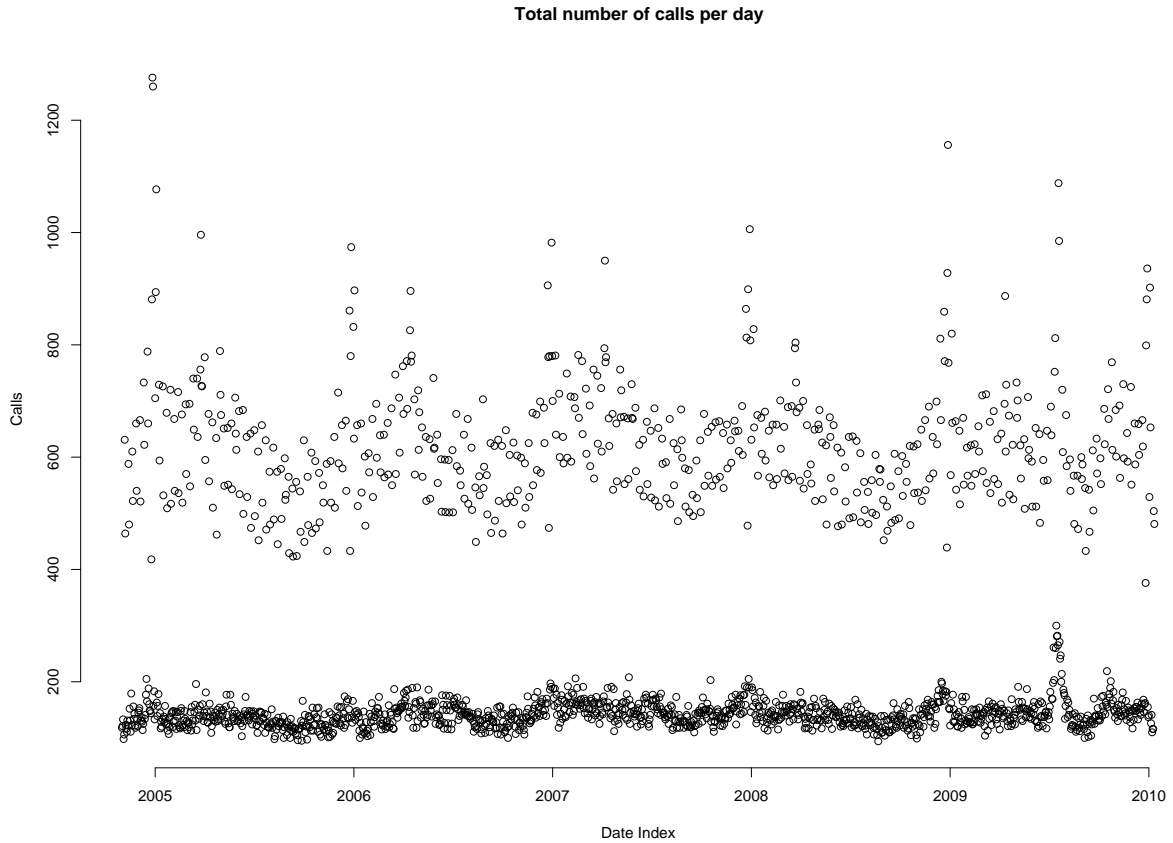


Figure 3.12: Calls per day between 01/11/2004 and 10/01/2010

3.7 Swine flu

By February 2009 the *CPD* model was used to predict calls per day routinely, with updates every 2-4 weeks using new observations, and in practice this worked quite well.

At the end of April 2009 swine flu hit the press, and the move from containment to treatment of the virus in July 2009 meant NDUC saw a massive surge in unexpected calls. The National Pandemic Flu Service (NPFs) launched later that month, meaning patients had an alternative phonenumber and website to obtain the antivirals separate to their GP. Shortly after this launch NDUC's excess call volume started decreasing and activity levels were back to normal by August. The second, less pressured, wave occurred in October and November 2009, with the increase being generally on weekends rather than weekdays. This section states how equation 3.21 changed to include swine flu, with the detail of what happened during the pandemic discussed in Chapter 5.

Figure 3.12 shows calls per day between 01/11/2004 and 10/01/2010. Just past the middle of 2009 you can see an uncharacteristic and unprecedented spike in activity, particularly in weekdays, and similar to behaviour we would associate with Christmas. This is also illustrated in figure 3.13, which shows the same calls per day between May and December 2009. Two key things happened here:

1. Movement from containment to treatment on 02/07/2009, and
2. Dedicated swine flu website and phonenumber to receive antivirals went live 23/07/2010.

Recall equation 3.21. For NDUC to continue to use the *CPD* model in the future we need

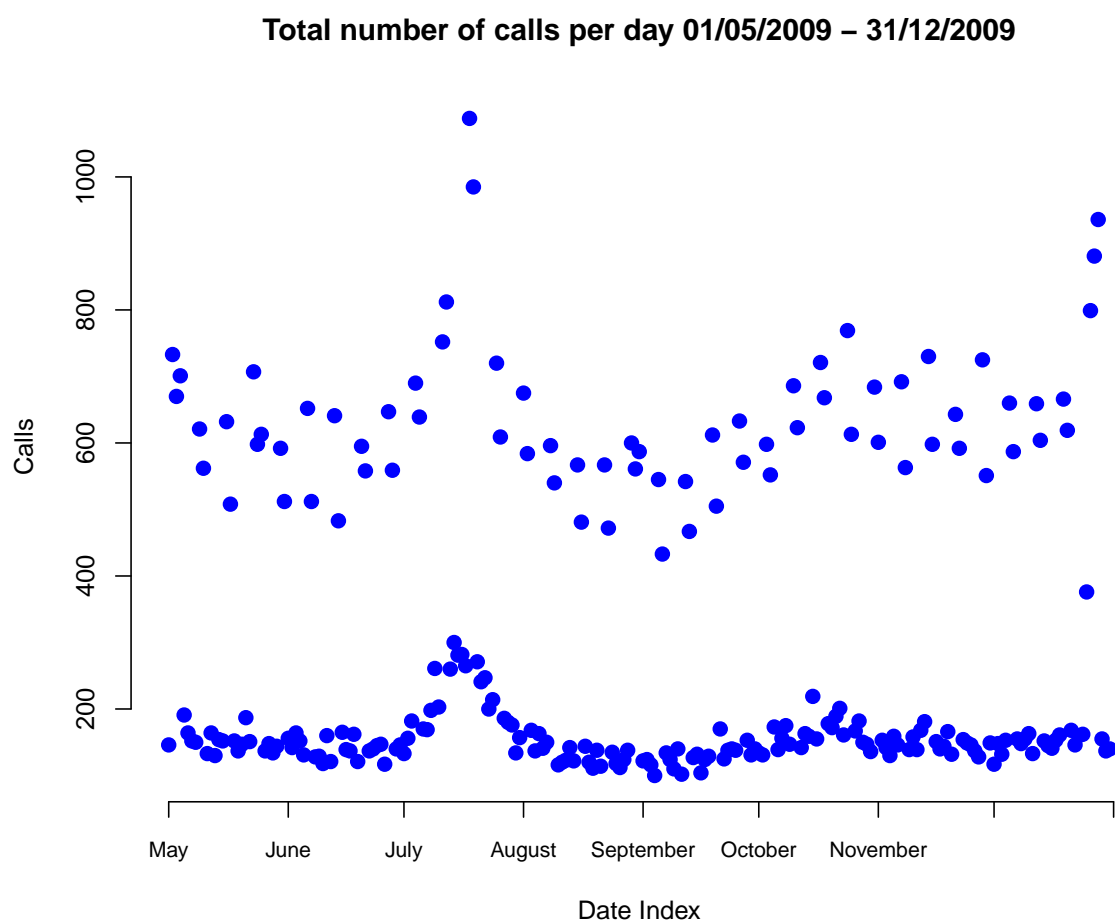


Figure 3.13: Calls per day between 01/05/2009 and 31/12/2009

to identify special dates with surges in activity. Consider equation 3.28 and define two new dummy variables:

- *sflu*: any day in the two swine flu waves 04/07/2009 - 02/08/2009 (inclusive), and 17/10/2009 - 27/10/2009 (inclusive).
- *sfluline*: any day when the NPFS was running between 23/07/2009 - 11/02/2010 (inclusive).

$$\begin{aligned}
CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
& + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
& + \beta_9bhsat + \beta_{10}bh sun + \beta_{11}bhmon + \\
& + \beta_{12}maybh1 + \\
& + \beta_{13}xmasbhwd + \beta_{14}xmasbhwe + \\
& + \beta_{15}nydwd + \beta_{16}xmaswd + \\
& + \beta_{17}xmaswe + \beta_{18}nydwe + \\
& + \beta_{19}xmas4day1 + \\
& + \beta_{20}month : we + \beta_{21}dow : xmas4day1 + \\
& + \beta_{22}month : year + \beta_{23}month : time + \\
& + \beta_{24}time : easterbh + \\
& + \beta_{25}sflu + \beta_{26}sfluline + \\
& + \beta_{27}sflu : sfluline + \epsilon
\end{aligned} \tag{3.28}$$

In the R output below, we favour equation 3.28 over equation 3.21 during the period 01/11/2004 - 01/09/2010. However, consider figures 3.14 and 3.15 which show the standardised residual plots for equations 3.28 and 3.21, respectively, with pandemic periods coloured in pink. There is not much difference between the two, and both fail to predict well over the pandemic period. Furthermore, in times when there is no pandemic (for the foreseeable future), and since the NPFS stopped running in February 2010, terms *sflu*, *sfluline*, and, *sflu:sfluline* will be zero, and the *CPD* model in equation 3.28 will be the same as equation 3.21. Is equation 3.28 really superior? On the other hand, when forecasting for July 2011, say, if we use equation 3.21 this is likely to inflate the estimates, as we know July 2009 observations were not representative of normal activity. Therefore, in the absence of suitable alternatives equation 3.28 is used for forecasting after July 2009.

One possibility for further work would be to find another way of dealing with July 2009 observations. Perhaps Weighted Least Squares could be imposed, or a back forecasting method (similar to observations in June 2006) would be a better solution. For the purposes of this thesis, we continued to use equation 3.28, but recognise that this choice of method is not perfect.

Start: AIC=239.14

```

CPD ~ dow + month + year + time + easterfri + eastersat + eastersun +
      eastermon + bhsat + bhsun + bhmon + maybh1 + xmasbhwd + xmasbhwe +
      xmaswd + nydwd + nydwe + xmaswe + xmas4day1 + we:month +
      dow:xmas4day1 + month:year + time:easterbh + time:month

```

	Df	Sum of Sq	RSS	Cp
+ sflu	1	47319	1795352	186.43
+ sfluline	1	17679	1824992	220.70

```
<none> 1842671 239.14
```

```
Step: AIC=186.43
```

```
CPD ~ dow + month + year + time + easterfri + eastersat + eastersun +
      eastermon + bhsat + bhsun + bhmon + maybh1 + xmasbhwd + xmasbhwe +
      xmaswd + nydwd + nydwe + xmaswe + xmas4day1 + sflu + month:we +
      dow:xmas4day1 + month:year + time:easterbh + month:time
```

	Df	Sum of Sq	RSS	Cp
+ sfluline	1	20933	1774419	164.24
<none>			1795352	186.43

```
Step: AIC=164.24
```

```
calls ~ dow + month + year + time + easterfri + eastersat + eastersun +
      eastermon + bhsat + bhsun + bhmon + maybh1 + xmasbhwd + xmasbhwe +
      xmaswd + nydwd + nydwe + xmaswe + xmas4day1 + sflu + sfluline +
      month:we + dow:xmas4day1 + month:year + time:easterbh + month:time
```

	Df	Sum of Sq	RSS	Cp
+ sflu:sfluline	1	37402	1737017	123.00
<none>			1774419	164.24

```
Step: AIC=123
```

```
CPD ~ dow + month + year + time + easterfri + eastersat + eastersun +
      eastermon + bhsat + bhsun + bhmon + maybh1 + xmasbhwd + xmasbhwe +
      xmaswd + nydwd + nydwe + xmaswe + xmas4day1 + sflu + sfluline +
      month:we + dow:xmas4day1 + month:year + time:easterbh + month:time +
      sflu:sfluline
```

In conclusion, we find the *CPD* model is inadequate for predicting pandemic periods, and must seek other models. The model used to predict calls per day in the future remains as equation 3.21. An alternative method for forecasting during pandemics, through moving averages, is discussed in detail in chapter 5.

3.8 Validating the model over time

3.8.1 Changes in *CPD* over time

To conclude this chapter we will briefly comment on the change to *CPD* predictions over time. To do this consider two datasets and measure how they cope in predicting *CPD* for a week in February 2010.

1. Dataset 1: Calls per day between 01/11/2004 - 01/12/2009
2. Dataset 2: Calls per day between 01/11/2004 - 01/09/2010

We use equation 3.21 as the linear model for forecasting. Dataset 1 has no observations for February 2010, while Dataset 2 does. Predictions for February 2010 at the start of December 2009 and September 2010 are given in table 3.3, alongside the observations for the first week in February 2010. The fits do not change much over time, and the prediction intervals become smaller. The first week in February 2010 observed 1852 calls so the predictions made at the start of December were only out by 38 calls for the week.

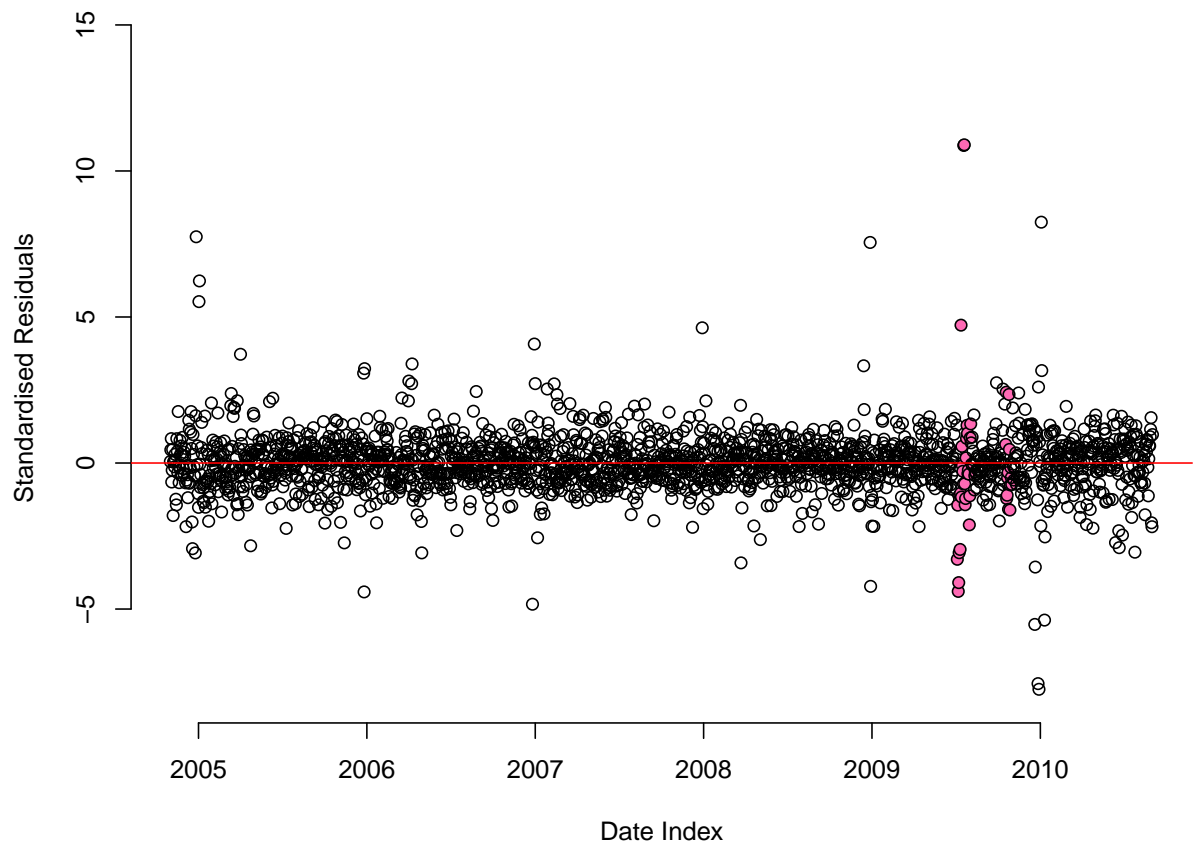


Figure 3.14: Standardised residual plot for equation 3.28 between 01/11/2004 and 01/09/2010

Day	Predicting in December 2009	Predicting in September 2010	Observation February 2010
Monday	125 (20,230)	128 (49,208)	133
Tuesday	116 (10,221)	119 (40,199)	126
Wednesday	117 (11,222)	120 (41,199)	118
Thursday	123 (17,228)	127 (48,206)	139
Friday	118 (12,224)	123 (44,202)	119
Saturday	660 (553,766)	660 (580,740)	671
Sunday	555 (449,662)	557 (477,637)	546
Total	1814	1834	1852

Table 3.3: Forecasting *CPD* for February 2010 in December 2009 and September 2010

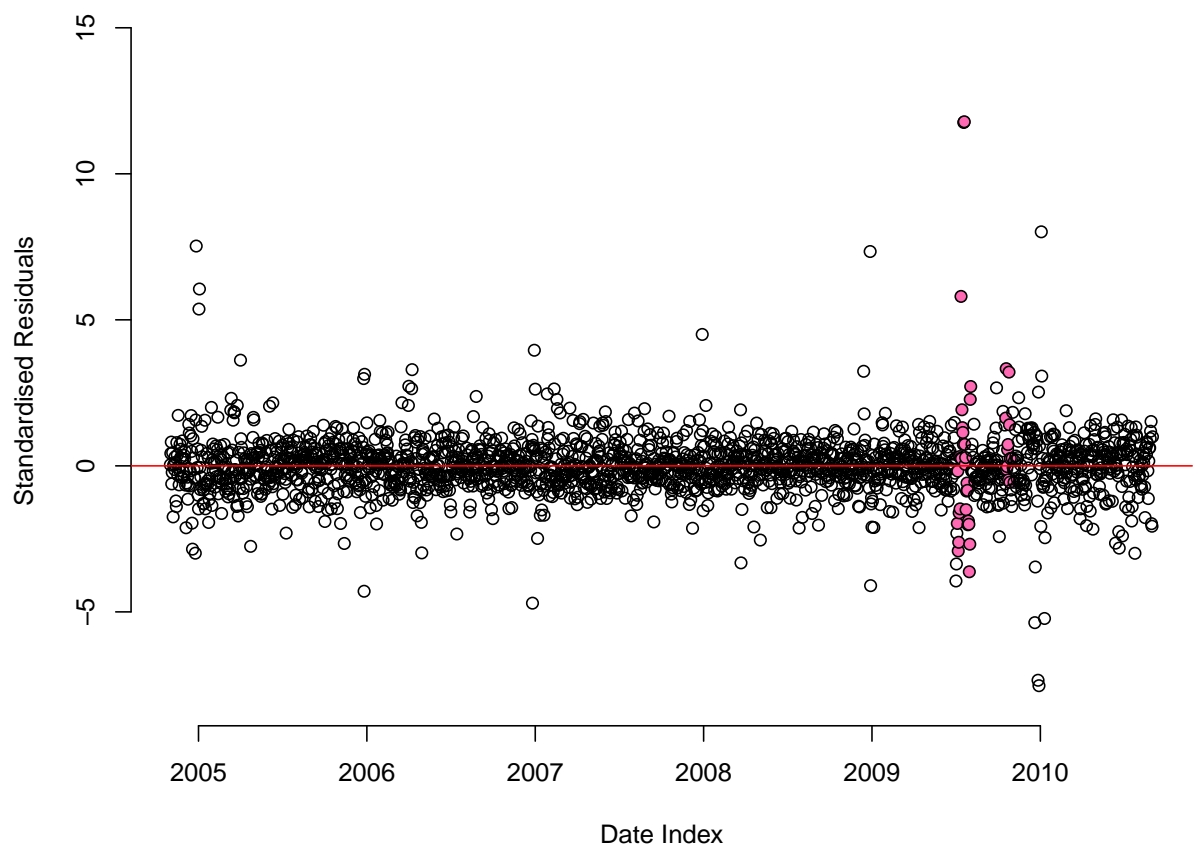


Figure 3.15: Standardised residual plot for equation 3.21 between 01/11/2004 and 01/09/2010

3.8.2 Periodically updating *CPD*

Every Monday morning NDUC holds an exception meeting where the performance and activity for the previous week is discussed. Here, a short comment will be made about how well the *CPD* forecasts matched the observed activity. It is from this monitoring that staff can choose when to periodically update the *CPD* model.

If, after two weeks, it appears that *CPD* is not forecasting as well as it should, update the model with the new observations. If it is still not fitting well, this could be due to external reasons, such as a pandemic period, or after a publicity campaign, and alternative reasons should be sought for this.

If the model is fitting well, it is recommended to update it after 4 weeks. This means that there is nearly a whole new month added to the observations, capturing `month` further, which is an important covariate in equation 3.21.

3.9 Conclusions and Recommendations

The motivation behind the *CPD* forecasting model in equation 3.28 is well documented in this chapter, and the model itself has been critically evaluated. Applying linear regression in this case worked well in practice. *CPD* has been formally forecast in this way since February 2009 and has been communicated to NDUC's commissioners, improving NDUC's reputation, and allowing them to become a reliable source of data information. It should not be forgotten that in exceptional times of activity, like Christmas and pandemics, this model does not work well, and alternative methods should be sought.

Chapter 4

Modelling Call Centre Demand

This chapter details the techniques implemented at NDUC to forecast intra-day call demand.

4.1 Establishing the Daily Profile

After establishing the forecasting procedure for calls per day, *CPD*, intra-day forecasting was necessary to determine staff levels and build rotas. It was clear that call arrivals into the centre were not constant, and instead there were peaks and troughs in demand, depending on the time of day. This led to the growth of the *Daily Profile* model which answered the following question; for any day in the future, how many calls will arrive in any 30 minute period?

4.1.1 Data Collection

We use a version of the dataset discussed in Chapter 3; every out-of-hours case from January 1st 2005, taken from the second run of the data extraction (recall figure 3.1 which highlighted an error in the first data extract), but this time ending in September 2010, rather than June 2009. Between 01/01/2005 and 01/09/2010 (inclusive) there were a total of 594,980 individual out-of-hours observations over the various days, times and seasons. The dataset is defined exactly as in Chapter 3, with various dummy variables indicating particular characteristics of any day.

The first step in this process was to discretize the data into 30 minute time slices and define the time periods to forecast over. For operational reasons, we define the forecasting periods by shifts:

Monday	= Monday 18:00:00 until Tuesday 07:59:59
Tuesday	= Tuesday 18:00:00 until Wednesday 07:59:59
...	
Friday	= Friday 18:00:00 until Saturday 07:59:59
Saturday	= Saturday 08:00:00 until Sunday 07:59:59
Sunday	= Sunday 08:00:00 until Monday 07:59:59

Monday, Tuesday, ..., Friday belong to weekday daily profiles, which have 28 unique time periods. Saturday and Sunday belong to weekend daily profiles, which have 48 unique time periods. As with the *CPD* model, we separate for different days of week, months, and holidays. For example, Good Friday will only have previous Good Friday observations, and Good Fridays are removed from normal Fridays in March or April.

To illustrate the application, consider the subset ‘Saturdays in August’. Saturdays in August are defined as any Saturday in August in the dataset that does not precede an August Bank Holiday Monday. Figure 4.1 shows observations for Saturdays in August, with the dates of these observations detailed in table 4.1. These observations are used to forecast for Saturdays

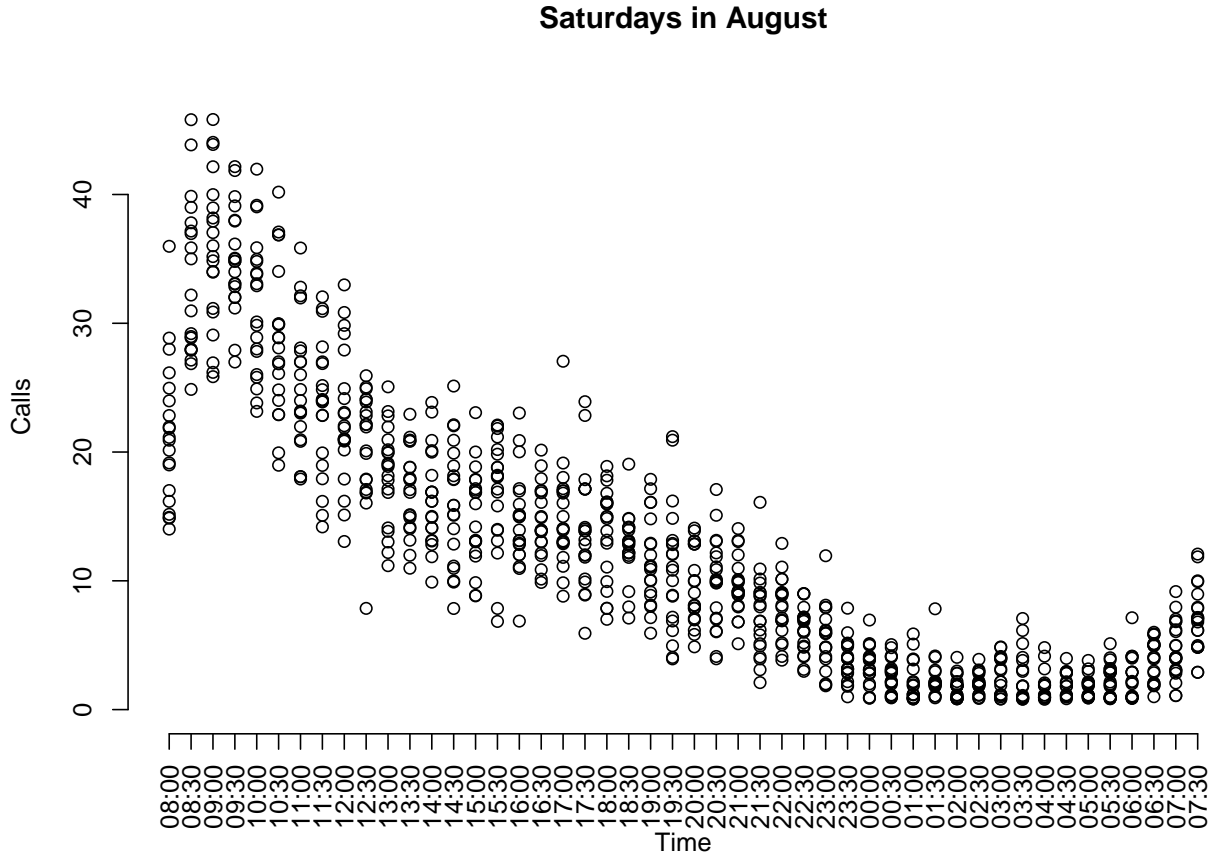


Figure 4.1: Previous half-hourly observations with random variation

in August 2011. There are a total of 960 observations; 20 observations for each of the 48 time points. There is a clear pattern with a sharp increase between 08:00 and 09:29, a steady decrease between 09:30 and 12:59, a slower decrease between 13:00 and 23:59, and a quieter period overnight between 00:00 and 06:59.

Note also that there is quite significant variation in the number of calls arriving into each time slot, and this variation does not seem to be constant.

4.1.2 Locally Weighted Regression

Locally weighted regression is a multivariate smoothing procedure which fits a regression surface to a set of data points [11] [12]. Let y_i , $i = 1, \dots, n$ be measurements of a response variable and let $x_i = (x_{i1}, \dots, x_{ip})$, $i = 1, \dots, n$ be the corresponding vector of measurements of p predictors. The relationship between the response and predictors is

$$y_i = g(x_i) + \epsilon_i,$$

where we assume $\epsilon_i \sim N(0, \sigma^2)$ are random errors and g is a smooth function of the predictors.

Locally weighted regression provides an estimate $\hat{g}(x)$ at each value x_1, \dots, x_p . The estimate of g at x_j , ($j = 1, \dots, p$) is obtained from a neighbourhood of points, weighted according to their distance from x_j . Points close to x_j have a large weight, and points far away from x_j have a small weight.

Year	Dates
2005	6 th , 13 th , 20 th
2006	5 th , 12 th , 19 th
2007	4 th , 11 th , 18 th
2008	2 nd , 9 th , 16 th , 30 th
2009	1 st , 8 th , 15 th , 22 nd
2010	7 th , 14 th , 21 st

Table 4.1: Dates in August used for the production of forecasts for ‘Saturdays in August’

To carry out locally weighted regression we need a distance function, ρ , a weight function, W , and the specification of a neighbourhood size. Further details of Cleveland’s loess concept are edited and well compiled in Chapter 8 of Chambers and Hastie’s book [10].

In our case we smooth by using the `loess` function in the stats package for R [20]. Here, fitting is done locally and the size of the neighbourhood is controlled by the argument `span`. Due to the curvature illustrated in figure 4.1 we use locally quadratic fitting by least squares. This is controlled by the arguments `degree = 2`, and `family = ‘gaussian’`, which are defaults for `loess` in R.

The variation of the overall pattern shows some, but not much, unusual behaviour. The observations at each time period seem evenly spread with possible exceptions around 12:30 and 17:00. Therefore we begin with the assumption that errors are Gaussian. In what follows `span` is always less than 1, so the neighbourhood includes a proportion, `span`, of the points, and these have tricubic weighting W .

As previously mentioned, the subset of Saturdays in August has 960 observations of two variables; the response variable which is calls, which we will denote by C (arranged in a $p \times n$ matrix), and the predictor is the half-hourly time period T (arranged in a $p \times 1$ matrix). Create a `data.frame` in R so the observations are arranged in a $p \times (n + 1)$ matrix (where $p=48$ and $n=20$) like so:

$$A_{p,n+1} = \begin{pmatrix} t_1 & c_{1,1} & \cdots & c_{1,n} \\ t_2 & c_{2,1} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ t_p & c_{p,1} & \cdots & c_{p,n} \end{pmatrix}$$

where $c_i = (c_{1,i}, \dots, c_{48,i})$, $i = 1, \dots, 20$, are the previous observations for day i , (the response), and $t = (t_1, \dots, t_{48})$ are the 30 minute time periods running from 08:00 on the Saturday until 07:59 the next Sunday. We can create the loess modelling in R by using the R command:

```
> loess(C~T,span=0.15)
```

A span of 0.15 was chosen because it worked well in practice, as illustrated in figure 4.2. Anything higher than this underestimated the 08:00 peak (figure 4.3), and anything more becomes too spiked (figure 4.4).

A span of 0.15 seemed to work for other days, with Tuesdays and Sundays in August loess fit illustrated in figures 4.5 and 4.6, respectively. Choosing a span of 0.15 for weekdays (14 hour periods) and weekends (24 hour periods) means the smoothing will be different for weekdays and weekends, and this can be seen by comparing the smoothness of figures 4.5 and 4.6. However, in practice this choice worked well to capture arrivals, and was adopted in the organisation.

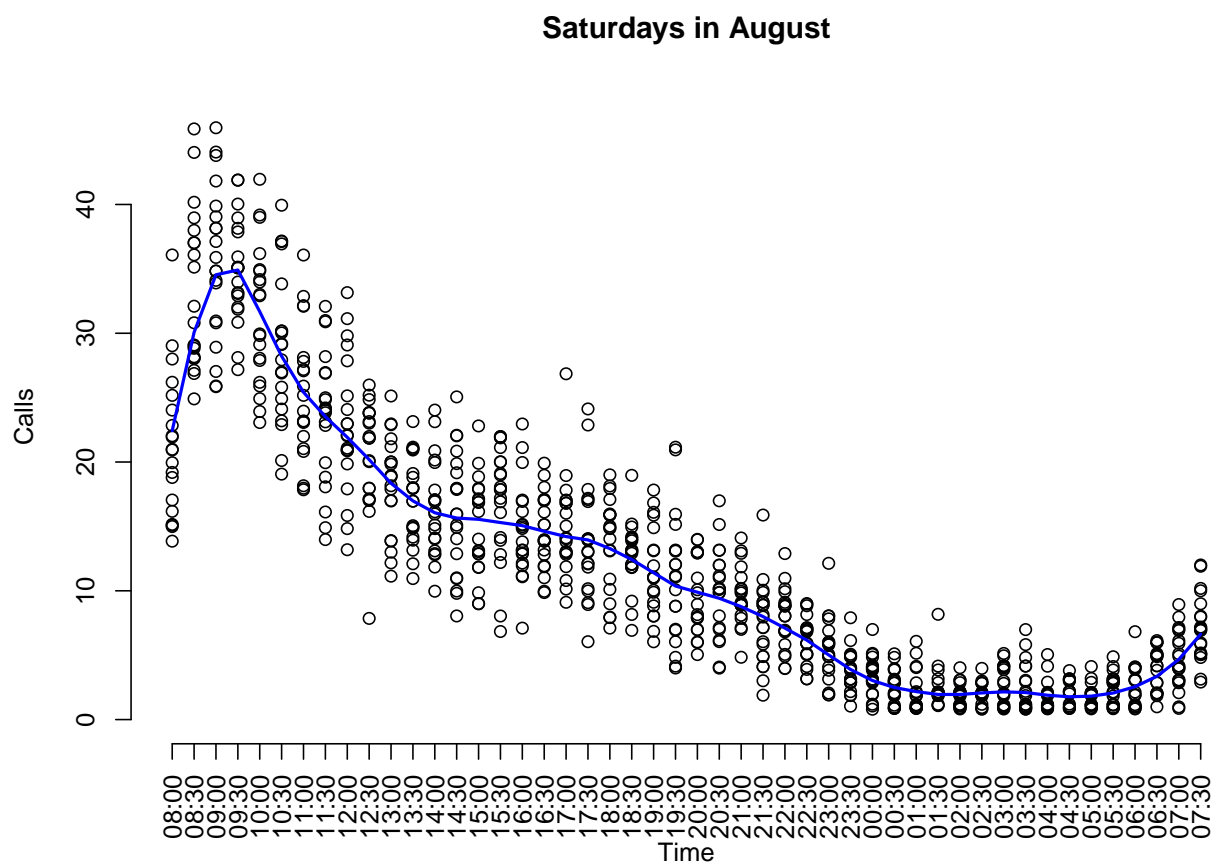


Figure 4.2: Previous half-hourly observations with random variation with loess fit using $\text{span}=0.15$

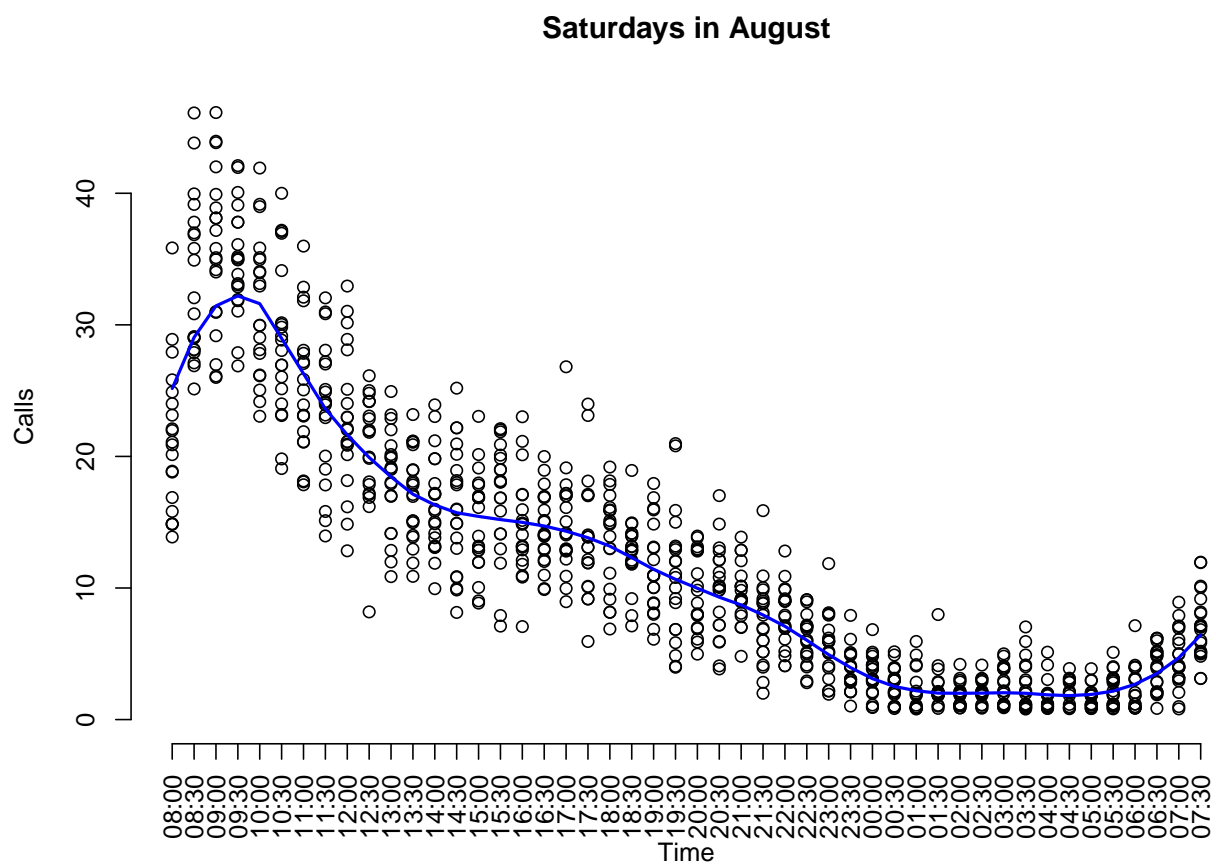


Figure 4.3: Previous half-hourly observations with random variation with loess fit using $\text{span}=0.25$

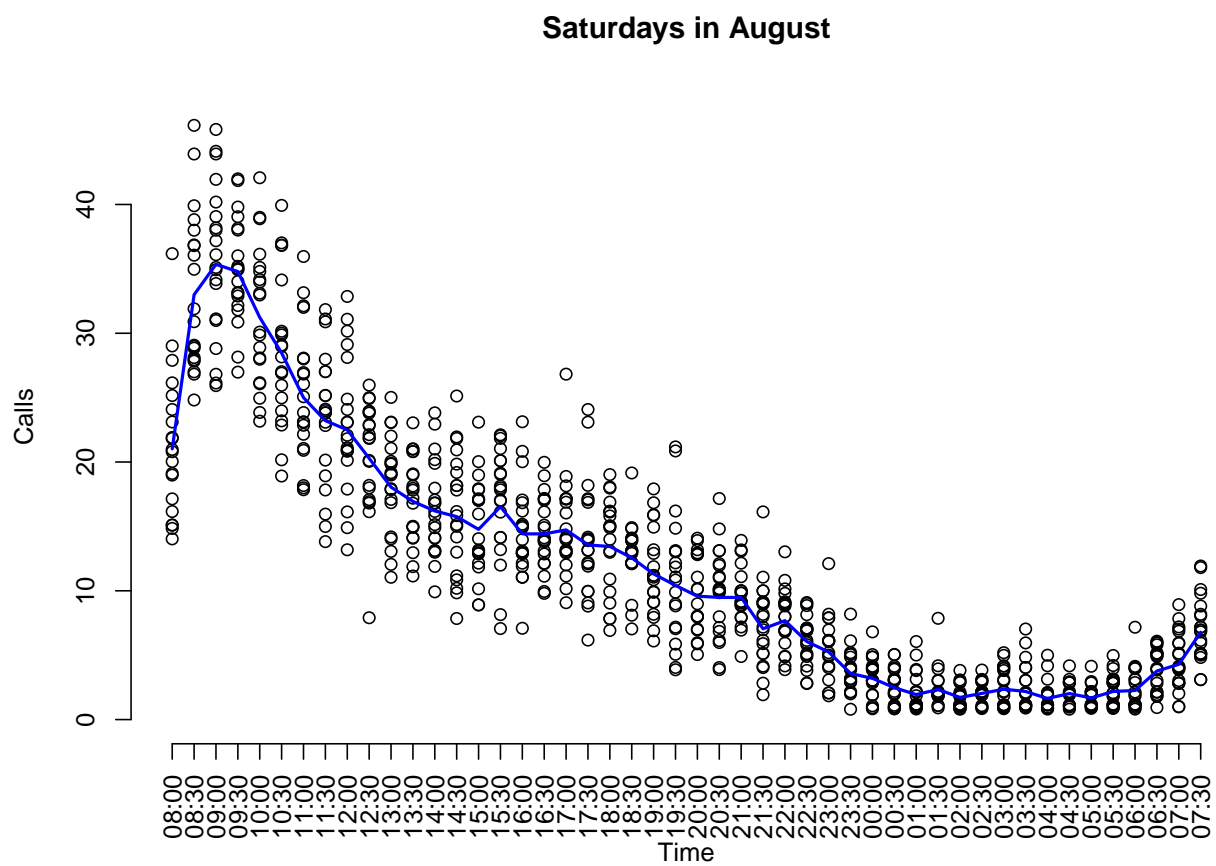


Figure 4.4: Previous half-hourly observations with random variation with loess fit using $\text{span}=0.05$

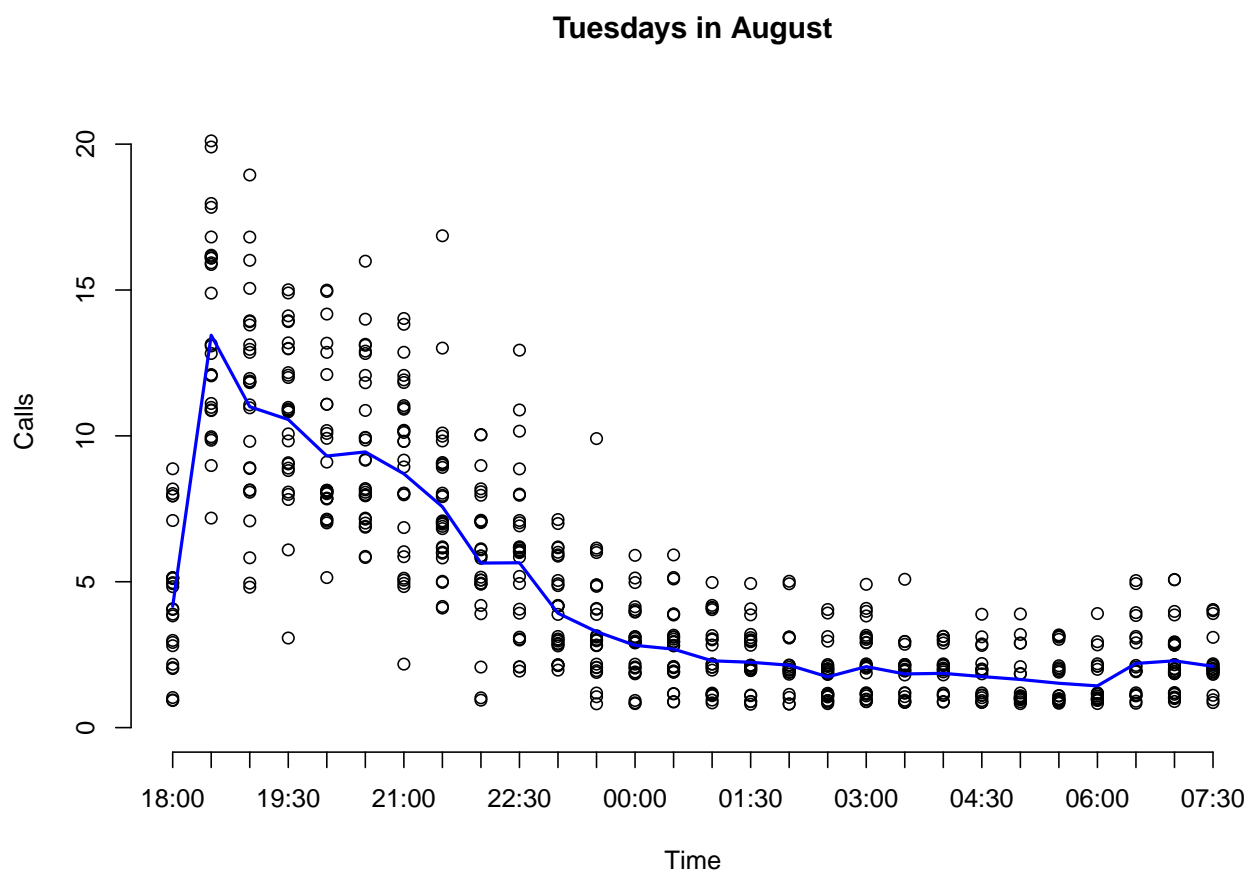


Figure 4.5: Tuesdays in August: half-hourly observations with loess fit using span=0.15

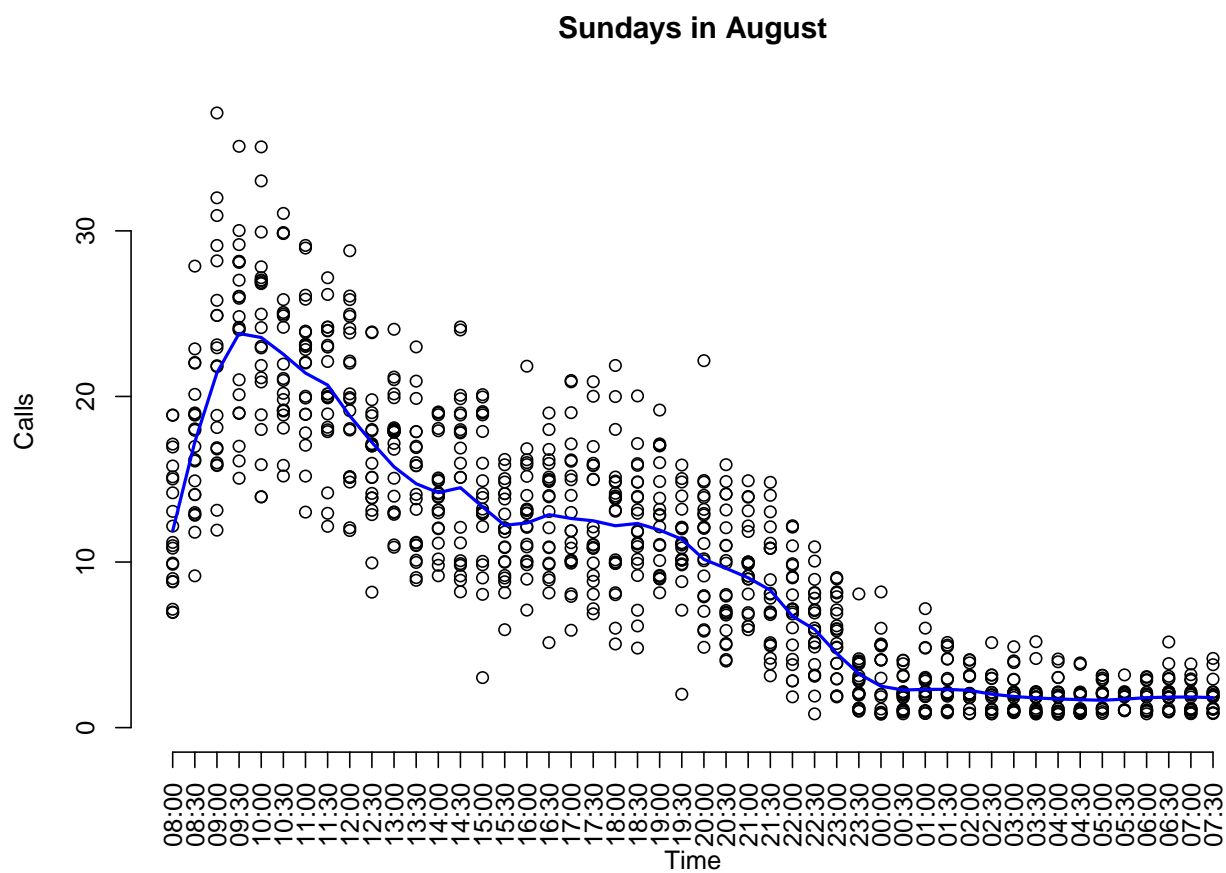


Figure 4.6: Sundays in August: half-hourly observations with loess fit using $\text{span}=0.05$

4.2 Establishing the Prediction Error

After establishing a reasonable method for predicting future patterns of call arrivals, we now need to place confidence bounds on these predictions. This gives NDUC an accurate view on possible maximum and minimum arrival rates. This knowledge will be essential for maximum capacity planning around pressure periods like Christmas. During the busiest times, how many phonelines will be occupied? Do they need more phonelines to cope with demand?

Recall the relationship between the response and predictor; $y_i = g(x_i) + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma^2)$, σ is unknown, and g is a smooth function. In least squares regression, when σ is unknown, we can generate standard errors of the fitted values from equation 4.1 [14], where $\hat{\sigma}$ is the estimate of σ and $s_{xx} = \sum_i^n (x_i - \bar{x})^2$. These standard errors are used when generating confidence intervals on observations.

$$sefit = \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{s_{xx}}} \quad (4.1)$$

Furthermore, when we are trying to predict a new observation $y_0 = g(x_0) + \epsilon_0$, which is independent of the observed data, the associated prediction error is stated in equation 4.2 [14].

$$sepred = \hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{s_{xx}}} \quad (4.2)$$

The confidence and prediction intervals are then generated by $\hat{y}_0 \pm Z_{1-\alpha/2} sefit$, and $\hat{y}_0 \pm Z_{1-\alpha/2} sepred$, respectively, where Z is the standard Normal distribution function, and α is the level of significance. It is worth noting that the fits when using confidence and prediction intervals are the same, but the prediction interval is wider than the confidence interval.

The least squares errors from equations 4.1 and 4.2 can be used to relate the fit and prediction errors by equation 4.3.

$$sepred^2 = \hat{\sigma}^2 + sefit^2 \quad (4.3)$$

This suggests that if we have a good estimate of σ^2 and a good estimate of $sefit$, we can derive an approximate prediction error. As it happens, standard loess packages give an estimate of σ^2 and of $sefit$, so we will use these to generate plausible $sepred$ for our prediction intervals.

We also make note here that we are using a Normal-based prediction interval, and this depends on a strong Normality assumption. In practice this may not be appropriate, however, in what follows, we assume it is true.

4.2.1 Investigating t errors

In Chapter 8 of Chambers and Hastie [10] there is coherent discussion on the application of confidence intervals using loess, but there is no comment on errors suitable for prediction. Therefore we decided to test the t-distribution for errors, as this is used in generating confidence and prediction intervals in the least squares case, when n is small.

The next thing to choose was the degrees of freedom. In linear regression we get an unbiased estimate of σ^2 from $RSS/(n-p)$ (recall section 3.3), where RSS is the residual sum of squares, n is the number of observations and p is the number of covariates. When using loess you generate an output called *equivalent number of parameters*, or **enp**, in R. Therefore $n - enp$ would be an appropriate number to use for the degrees of freedom. Since σ^2 is unknown, and we are predicting y_0 , the degrees of freedom could also be $n - 2$. Therefore a t-distribution with $n - \max(2, enp)$ degrees of freedom was chosen.

In practice this only worked well over ‘busy’ parts of the shift, and the upper limit was too generous on the overnight part of the shift.

Consider figure 4.7 which contains three graphs. Firstly, it shows the previous observations for Saturdays in August, with the blue line illustrating the loess fit, and the upper and lower green lines illustrating the t errors. From this you can see the upper limit seems to be generous on the overnight part of the shift, compared to the observations. Along the bottom there is a normal quantile plot of the residuals (left) and a scatter plot of the standardised residuals (right). There is curvature in the quantile plot, and the standardised residuals are not randomly scattered. This motivated changing the errors to something more appropriate.

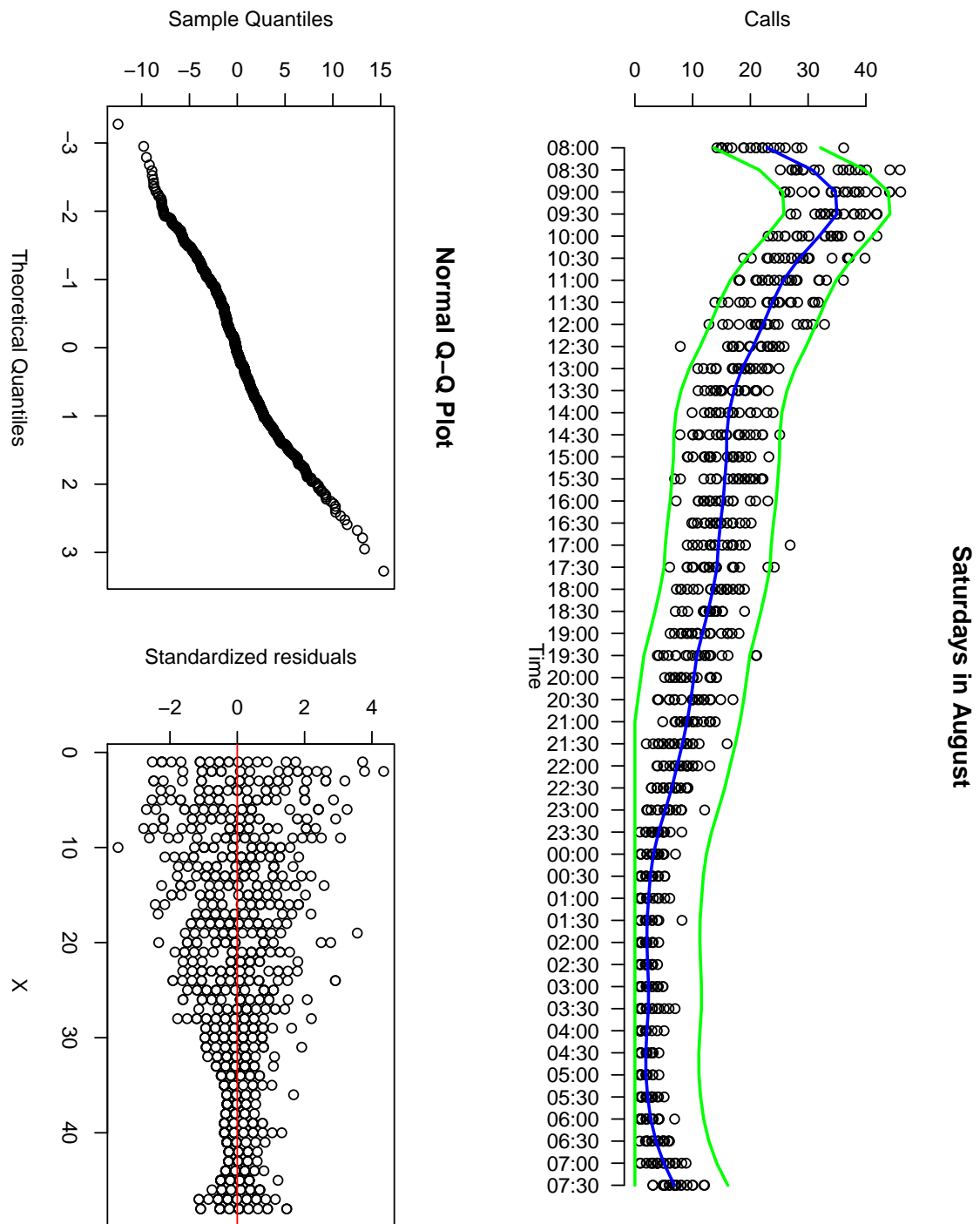


Figure 4.7: Previous half-hourly observations with random variation with loess fit and t errors

Time point	χ^2	df	p value	Time point	χ^2	df	p value
t_1	19.55633	13	0.10683870	t_{25}	17.89380	7	0.01245844*
t_2	26.33990	12	0.00960552*	t_{26}	18.56536	8	0.01736515*
t_3	24.29513	12	0.01853983*	t_{27}	12.68433	7	0.08018356
t_4	23.09211	10	0.01041164*	t_{28}	10.30718	9	0.3261944
t_5	22.63181	11	0.01991085*	t_{29}	9.62981	7	0.2105405
t_6	27.88999	11	0.003365318*	t_{30}	10.90561	5	0.05328381
t_7	25.47500	10	0.004514292*	t_{31}	16.20925	7	0.02327188*
t_8	24.65103	11	0.01025096*	t_{32}	5.93290	5	0.3128029
t_9	19.74197	13	0.1018335	t_{33}	7.89078	4	0.09566193
t_{10}	27.25200	8	0.000639516*	t_{34}	3.10819	4	0.5398868
t_{11}	15.04379	10	0.1304734	t_{35}	5.689893	5	0.3375722
t_{12}	19.08336	8	0.01442015*	t_{36}	6.130755	3	0.1054191
t_{13}	12.68039	10	0.2420946	t_{37}	10.04525	2	0.006587205*
t_{14}	18.84898	11	0.06385873	t_{38}	3.67406	3	0.2988793
t_{15}	19.53425	9	0.02101462*	t_{39}	6.32985	4	0.1758323
t_{16}	20.96129	10	0.0213651*	t_{40}	7.54552	6	0.2733249
t_{17}	18.59836	9	0.02883298*	t_{41}	4.63632	4	0.326688
t_{18}	9.80782	9	0.3662662	t_{42}	3.28285	3	0.3500369
t_{19}	17.65048	10	0.06115356	t_{43}	1.82016	3	0.6105588
t_{20}	28.48130	8	0.0003908762*	t_{44}	7.48336	3	0.05798758
t_{21}	14.75672	10	0.1411863	t_{45}	8.86022	4	0.06468998
t_{22}	24.15219	6	0.0004896407*	t_{46}	10.21111	4	0.03701775*
t_{23}	9.49610	10	0.4857554	t_{47}	5.04210	7	0.6548254
t_{24}	28.01161	10	0.001797541*	t_{48}	14.31237	6	0.02633493*

Table 4.2: Goodness-of-fit test for the Poisson distribution

4.3 Investigating Poisson errors

We explore the distribution of the number of arrivals $c_{1,i}, \dots, c_{48,i}$ during each time slice $t_j \in t$, $j = 1, \dots, 48$, using the function `goodfit` from the package `vcd` [13]. This fits a Poisson distribution for goodness-of-fit tests through maximum likelihood by specifying `type = "poisson"` and `method = "ML"` in the arguments. Using `summary(goodfit)` in R outputs the likelihood ratio statistic and p values. This is quoted in table 4.2. We fail to reject the hypothesis at the 5% level of significance for 27 of the time points. Therefore we will assume the that arrivals are approximately Poisson.

If the expected number of occurrences in this interval is $\lambda > 0$, then the probability that there are exactly x ($x = 0, 1, 2, \dots$) occurrences is equal to

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}. \quad (4.4)$$

The mean and variance of the Poisson distribution are calculated from equations 4.5 and 4.6 [14], respectively.

$$E[X] = \lambda \quad (4.5)$$

$$Var[X] = \lambda \quad (4.6)$$

A $\text{Poisson}(\lambda)$ distribution can be thought of as the sum of λ independent $\text{Poisson}(1)$ distributions. For sufficiently large values of λ , the Normal distribution with mean λ and variance λ approximates the Poisson distribution well [14]. Recall our assumptions about ϵ and $c_{i,j}$: $\epsilon_i \sim N(0, \sigma^2)$, and $c_{i,j} \sim \text{Pois}(\lambda)$. We transform by taking $\tilde{c}_{i,j} = \sqrt{c_{i,j}}$, and then apply loess to the transformed values to obtain the predicted value, \hat{t}_j , before transforming back. The estimated arrival rate for time slice t_j is now $\hat{\lambda}_j = \hat{t}_j^2$, and the $\text{Poisson}(\hat{\lambda}_j)$ is used to generate a forecast and a confidence envelope for the number of calls arriving into each time slice.

This provides a piecewise prediction interval for arrival rates for a specific day, as illustrated in figure 4.8. Figure 4.8 contains three graphs: previous observations with loess fit and Poisson errors overlaid, a Normal quantile plot of the residuals and a standardised residual plot. On the top graph the loess fit is blue, as before, and the Poisson errors are red. These errors look superior to the t errors, and fit the observations better. The quantile plot is quite straight, and the standardised residual plot has improved from figure 4.7. There does, however, appear to be pattern in the residuals on the overnight part of this shift, between 22:30 - 07:59, or $X = 30, \dots, 48$. Figure 4.9 shows three individual days of residuals in this time period on three separate lines, and between 03:30 - 05:59 ($X=39, \dots, 44$) something seems to be wrong. This indicates that the model may not be quite right, but the residuals remain small. The period 03:00 - 05:59 is a quiet part of the shift, with little pressure, so operationally this will not matter too much.

Even with this discrepancy, using Poisson errors was superior to using t errors, and we progressed with forecasting this way. In practice, the square root transform and Poisson errors gives very good results, with no substantial anomalies.

Intuitively, you would normalise these so that the total arrivals over a day match the *CPD* model. In this case, however, the *CPD* model forecasts calls per day over a time period 00:00:00 - 23:59:59, while the *Daily Profile* model forecasts from either 18:00:00 or 08:00:00 on that day until 07:59:59 the next day, so we do not normalise. The reason for this is because of organisational requirements. NDUC's monthly performance report records daily activity, and this is defined as the same time period used in *CPD*. It makes sense to forecast *CPD* over the time period, as this will highlight unusual activity to the PCTs clearly. This was particularly useful in July 2009 when we were receiving far more calls than expected. On the other hand, when planning rotas and building shifts it is better to forecast arrival rates over a shift.

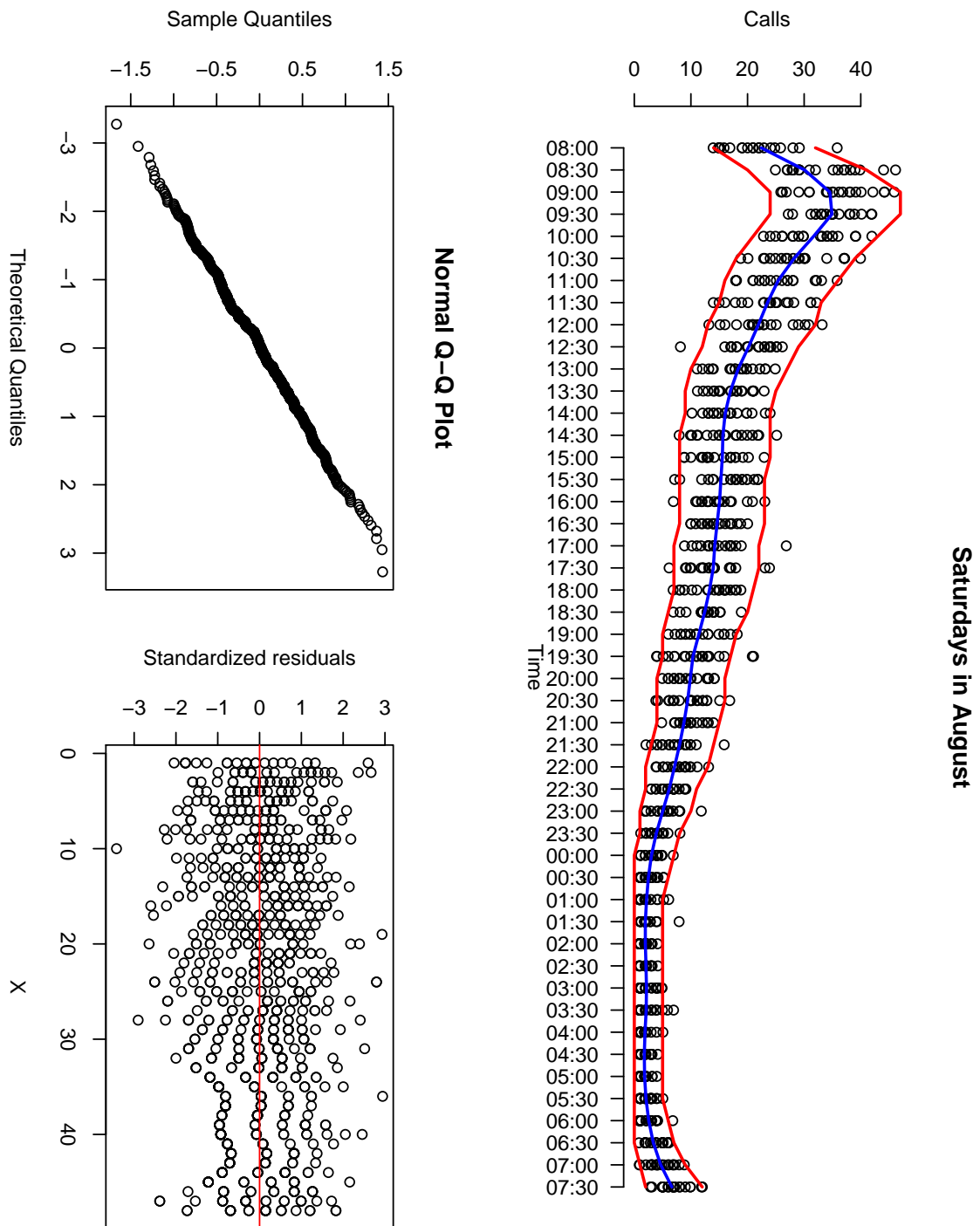


Figure 4.8: Previous half-hourly observations with random variation with loess fit and Poisson errors

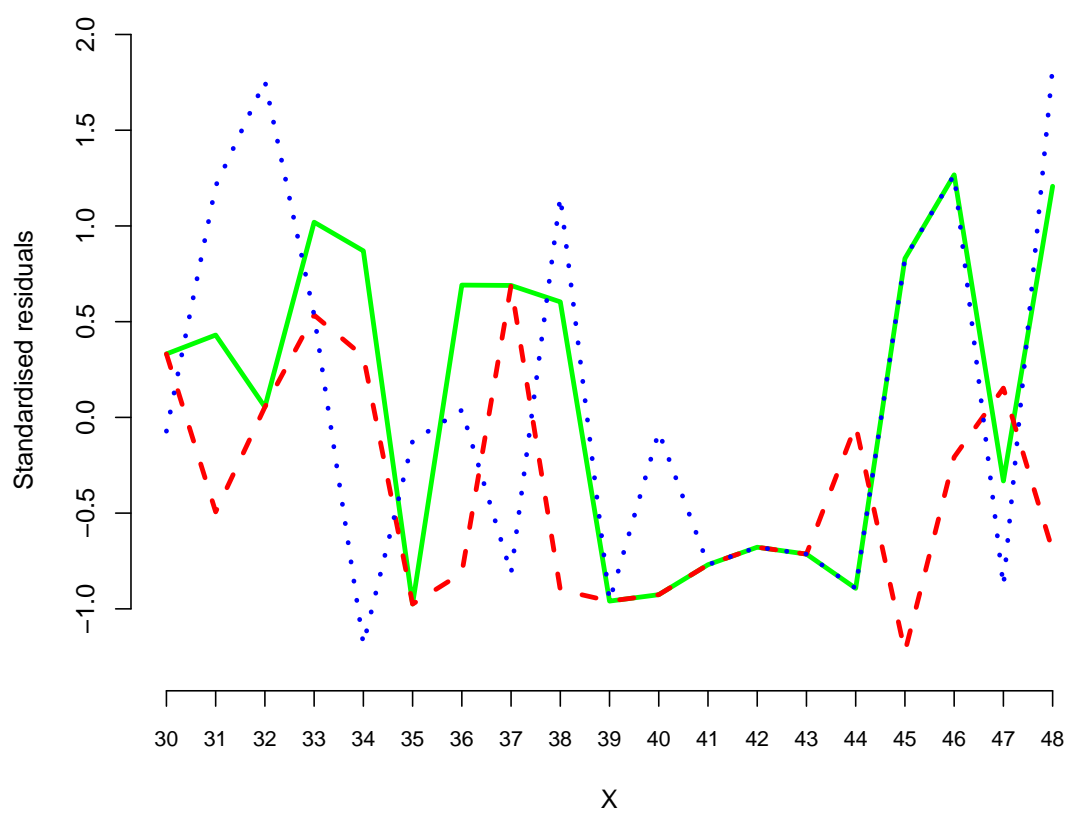


Figure 4.9: Checking the residuals: not random overnight

4.4 Monitoring the Daily Profile

4.4.1 Changes in *Daily profile* over time

The *Daily Profile* model has been running at NDUC for two years, and it's important to check how it changed over time. For consistency, consider again the subset 'Saturdays in August'.

The example previously described would be used to forecast arrival rates for Saturdays in August 2011. Suppose we consider two datasets:

- Dataset 1: Saturdays in August 2005 - 2008.
- Dataset 2: Saturdays in August 2005 - 2009.

Next, use the method described in this chapter to generate forecasts and prediction intervals to predict for Saturdays in August 2010.

Figures 4.10, and 4.11 show the fits, errors and observation plots, as well as the quantile and standardised residual plots for loess forecasting with Dataset 1 and Dataset 2 respectively. The main difference in having an extra year's observations is it smoothes the predictions and straightens the quantile plot slightly.

In practice, figures 4.12 and 4.13 show the predictions for Dataset 1 and Dataset 2, with the observations from the three Saturdays in August overlaid in green. Again, the only real difference is the smoothness of the fit and limits, but they both predicted well.

4.4.2 Periodically updating *Daily Profile*

Given the discussion in this chapter, particularly in section 4.4.1, it is recommended to update the *Daily Profile* model as soon as fresh data arrives. The way that is constructed, means you are able to forecast 30 minute arrivals eleven months in advance: August 2011 forecasts could be generated at the start of September 2010, and so on.

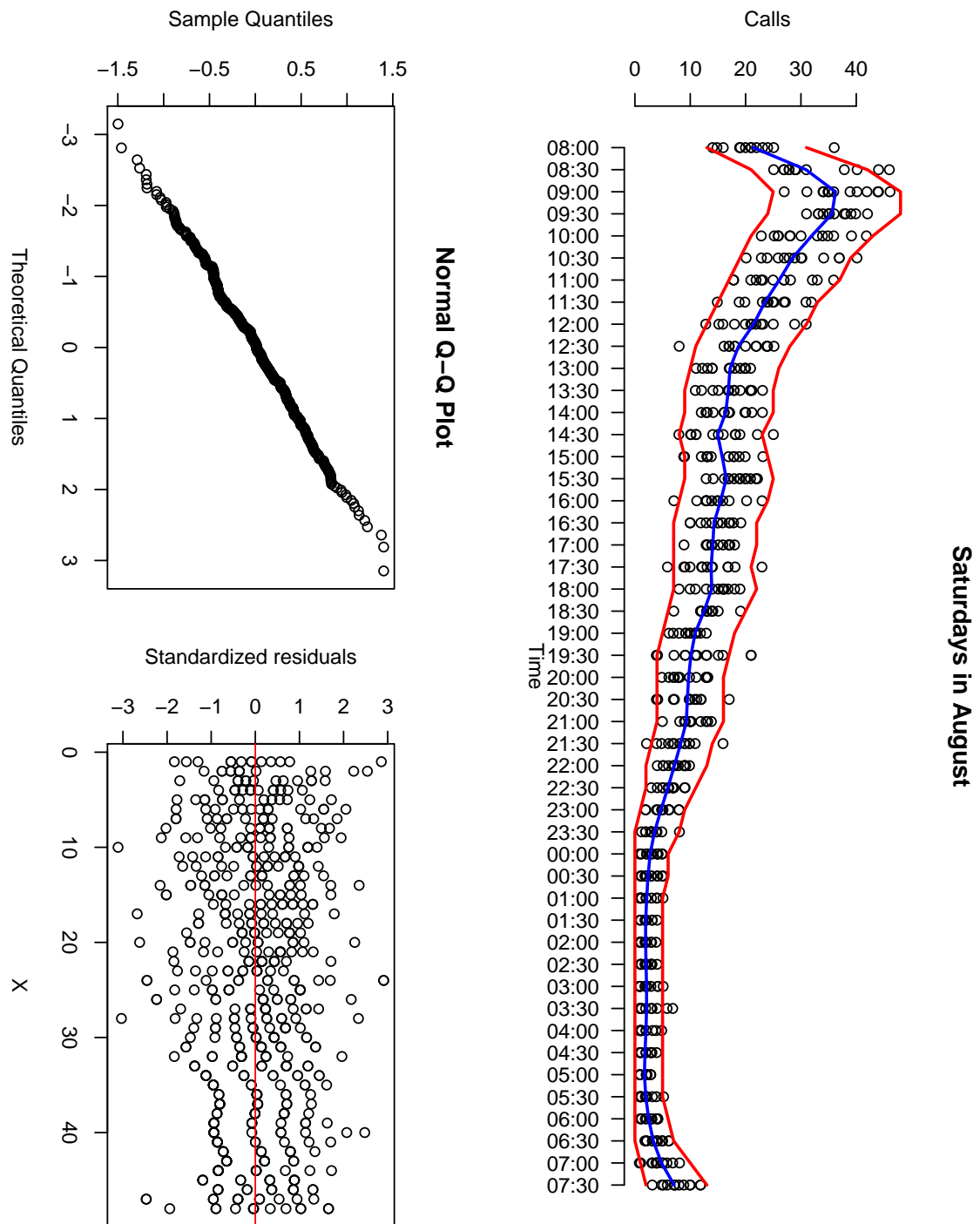


Figure 4.10: Previous half-hourly observations 2005-2008 with random variation with loess fit and Poisson errors

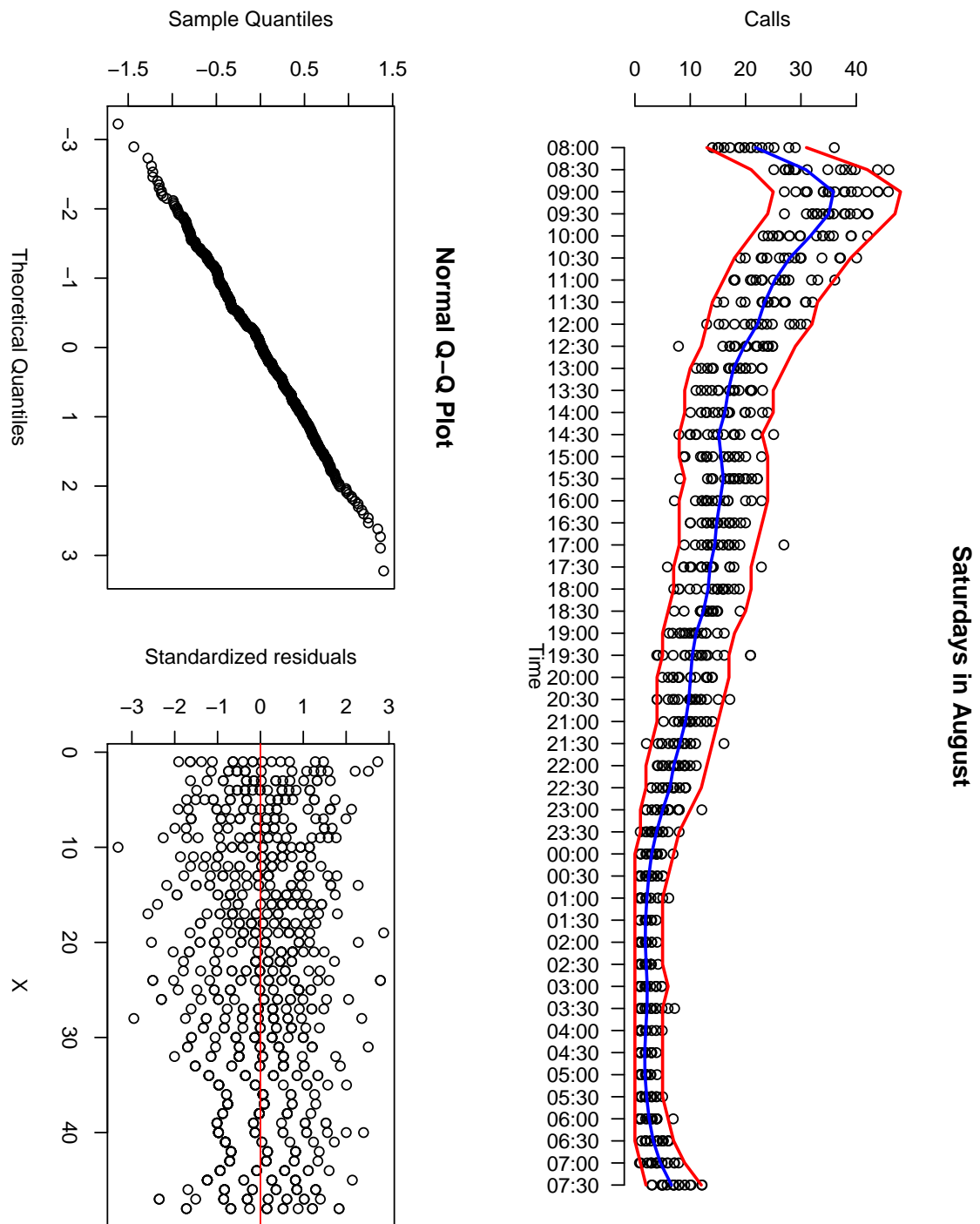


Figure 4.11: Previous half-hourly observations 2005-2009 with random variation with loess fit and Poisson errors

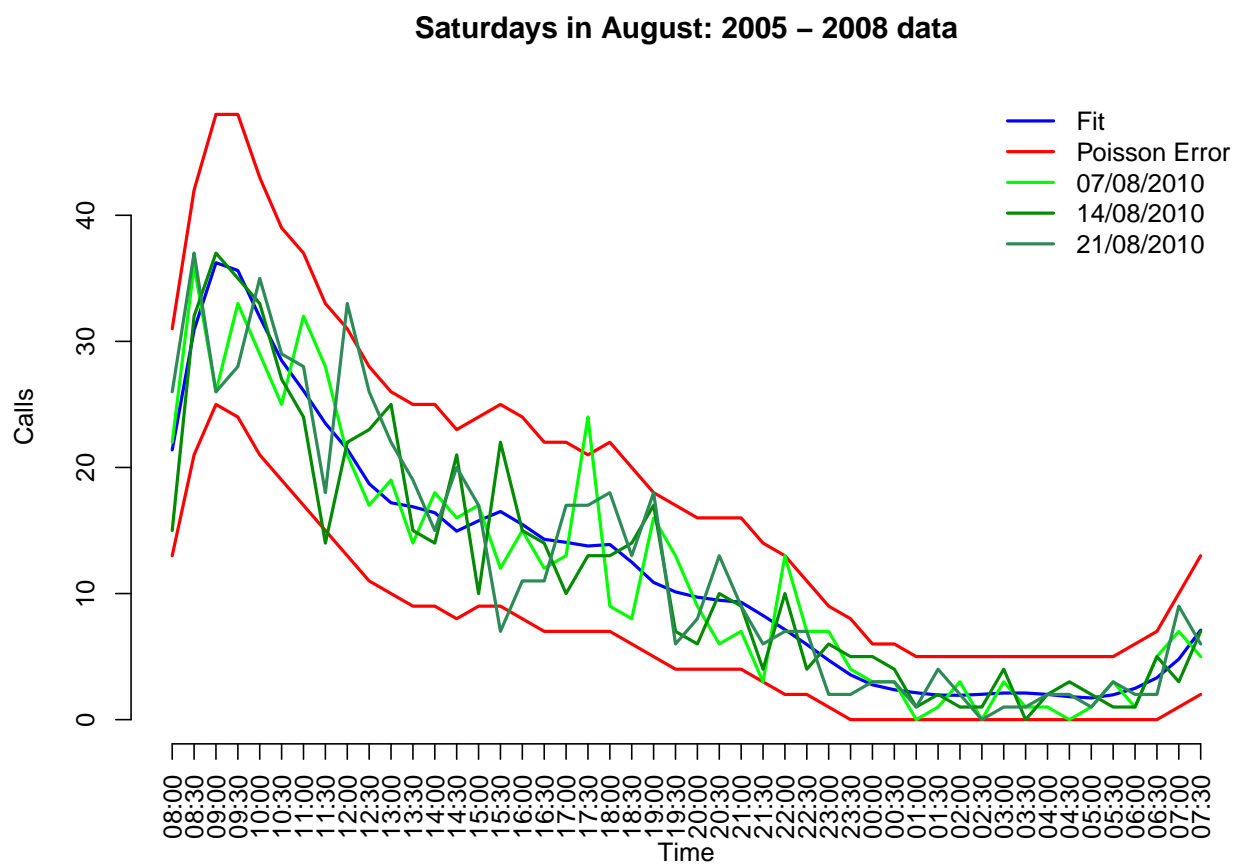


Figure 4.12: Loess fit and Poisson errors for 2005-2008 data with 2010 observations

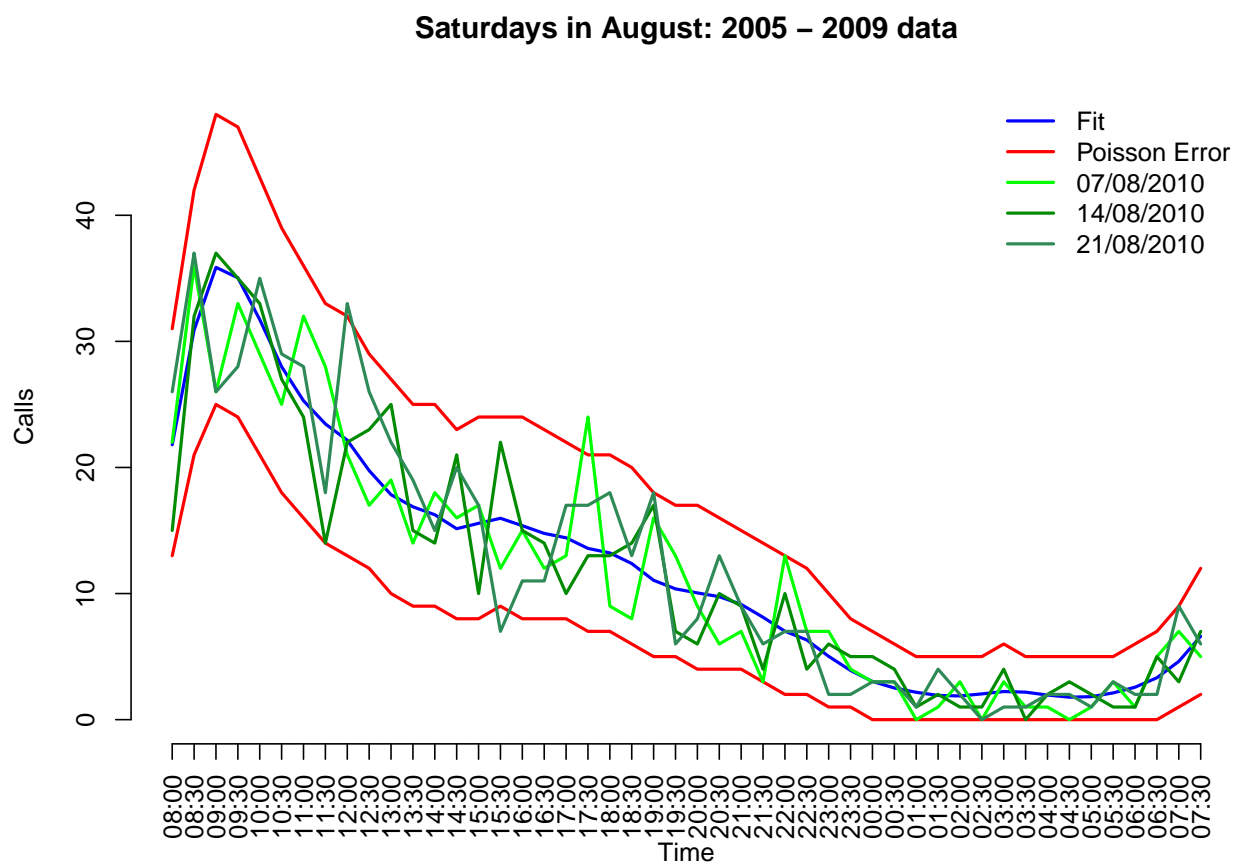


Figure 4.13: Loess fit and Poisson errors for 2005-2009 data with 2010 observations

4.5 Conclusions and Recommendations

In chapters 3 and 4 we created two separate models; *CPD* for forecasting calls per day, and *Daily Profile* for forecasting arrivals rates. There could be mistakes made by approaching forecasting in this way, for example, consider Property 3 from Avramidis *et al.*[7, pp 2]:

“There is strong positive association (correlation) between arrival counts in a time partition of a day”.

Applying this property, it may be worthwhile using *Daily Profile* to generate calls per day forecasts. Furthermore this may suggest updating your calls per day forecasts periodically during a shift is a better method, as discussed by Shen and Huang [21].

In practice, forecasting separately works for NDUC, however further work could be carried out to formalise the benefits and pitfalls of doing so. Currently, updating forecasts during shifts could not be applied in this case, as the data is harvested during ‘in-hours’ periods, ensuring there is no extra pressure on servers whilst patients are using the service.

To conclude; although there are some model faults, NDUC have two forecasting methods which produce good results. They have been used in several aspects of the business from workforce planning, to new bids for business. They improved the management information available to staff on all levels, and were particularly useful for commissioners during the swine flu pandemic in 2009.

Chapter 5

The Swine flu Pandemic

By February 2009 the *CPD* model was used to predict calls per day routinely, with updates every 2-4 weeks of new observations, and in practice this worked quite well. At the end of April 2009 swine flu hit the press in England. Widespread news coverage about the virus increased public awareness, and headlines such as the New Scientist's "Deadly new flu virus in US and Mexico may go pandemic" [2], caused public concern. Although the underlying messages tended to be to remain calm and not panic, the headlines did not support this message. An article published by the BBC in June 2009 read "First UK swine flu patient death" [3]. Upon further reading this patient had underlying health problems, and there was also a comment on the bottom from a Virologist entitled "Chances remote", advising the readers to not panic.

The uncertainty among the public was clear to us from GP feedback. Sessional GPs working at practices during the day recalled that some patients were confused about what to do in special cases. What about children, pregnant women, or those living in close quarters to someone who is going through chemotherapy?

When the NHS published the first guidelines for patients dealing with swine flu, they advised to contact your GP if you were experiencing symptoms, or had any concerns that you had the virus. Therefore, any patient following this advice in Northumberland, North Tyneside, Newcastle, or South Tyneside, out-of-hours would be directed to one of NDUC's GPs for advice. The initial publicity of the virus started just before the first May Bank Holiday 2009. During this holiday NDUC ran the GP out-of-hours service from 18:30:00 on Friday 1st May, until 07:59:59 the following Tuesday. Although it was clear there would only be a handful of genuine virus cases over this weekend, if any, we were still aware that anyone with normal flu symptoms, the worried-well, and those recently back from Mexico or America might call the service.

The national strategy for dealing with the pandemic developed and changed as the virus progressed. In May and June 2009 England remained in the containment phase. That meant, any suspected swine flu cases were isolated, swabbed and confirmed, and received treatment. On the 2nd July 2009, the UK moved from containment to treatment. This meant that anyone suspected of having swine flu received antivirals for treatment, without confirmation of the virus. It was after this date that we saw an unexpected influx of calls, described later in this chapter.

On 23rd July 2009, the National Pandemic Flu Service went live across England, and provided patients with antivirals for swine flu. By 30th July 2009 our call volumes seemed to be returning to normal. August and September 2009 remained at normal levels, but by October we were seeing unusually high activity again, particularly at weekends. The second wave of the pandemic hit England in October and November 2009, but without the extreme peak we saw in July.

This unusual activity between June and October 2009 is illustrated in figure 5.1, which shows the percentage increase above *CPD* forecasts between 01/06/2009 - 21/10/2009. We would normally see a small oscillation above and below the dashed grey line, and this wasn't

the case in July, or October. The vertical blue lines are when NDUC implemented and stood down its Rising Tide Policy, which is used in severe pressure periods, respectively.

All of these influences meant we needed to create an inflation term change the *CPD* model, and the following sections highlight the thoughts and methods used to tackle this problem.

5.1 The Health Protection Agency

The Health Protection Agency (HPA) is an organisation that supports the NHS, Department of Health (DoH), and other organisations by protecting UK public health mostly through advice. Throughout the swine flu pandemic they reported on swine flu incidences through updates on confirmed cases, press releases around travel advice, vaccines and treatment approaches, and then through Weekly Pandemic Flu Updates. Throughout this chapter I will refer to several of these reports, which are archived on the HPA website [4].

5.2 Strategic Health Authority

Strategic Health Authorities (SHAs) manage the local NHS on behalf of the secretary of state. There are ten SHAs across England, and NDUC falls under the North East group. They are an important link between the DoH and the NHS, and are responsible for making sure local health services are of high quality and performing well.

5.3 May Bank Holiday 2009

As previously alluded, in May 2009 the number of confirmed cases in the UK was low, and the impact on the out-of-hours service was likely to be those with the normal flu (which isn't too prevalent in May), and the worried well; like those returning from holidays in Mexico. We had little information on how the population would react to the media, but could look at previous service demand under normal conditions, and general characteristics of the population of the North East. Table 5.1 shows the monthly figures for previous Mays, and summing up the forecasts from the normal *CPD* model we expected the volume for May 2009 to be 10,058, busier than any previous May on record.

On 30th April 2009 the HPA gave an update of swine flu cases in England, which were around the London area, with one confirmed case in the North East of England. At this stage all cases had mild symptoms and were responding to treatment at home. Let's now consider the one confirmed case in the North East of England, which was associated with travel to Mexico. The HPA stated that close contacts to the infected could have antivirals as a precautionary measure. That is anyone who has been exposed to the confirmed case over the last week for more than one hour within a distance of 1 metre or less. The next day this news hit the press in the North East. Paul James, writing for The Journal [16], revealed that the confirmed case in the North East shared residence with two students at Newcastle University.

After consultation with sessional GPs and the operations team, it was agreed that a reasonable estimate for the excess calls over the Bank Holiday weekend was around 100, and an inflation coefficient I for *CPD* $I_1 = 5\%$ was implemented for the 2nd-4th May 2009, inclusive. Table 5.2 states how the forecasts looked in practice.

5.4 July 2009

The decision to move from containment to treatment of the virus was made on Thursday 2nd July 2009. This meant that patients who met certain criteria were able to obtain the antivirals, without confirming whether or not they had swine flu. These would be collected by a flu-friend

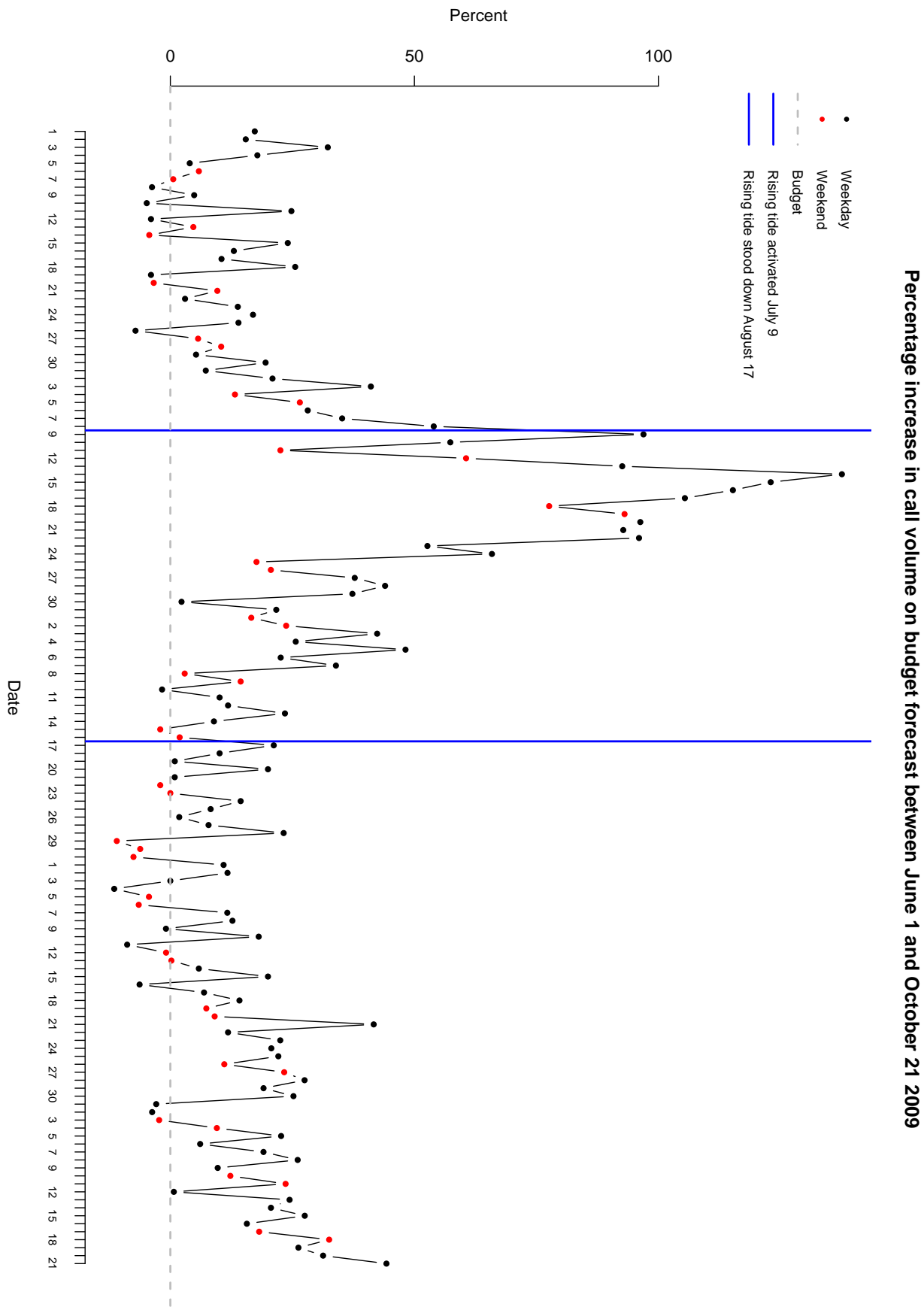


Figure 5.1: Percentage increase from *CPD* between 01/06/2009 - 21/10/2009

Table 5.1: Monthly figures for previous Mays

Year	Total cases over month
2005	9796
2006	9335
2007	9843
2008	9592

Table 5.2: Inflated fit, fit and actual *CPD* for May Bank Holiday 2009

Date	02/05/2009	03/05/2009	04/05/2009	Total
Normal fit	721	650	644	2015
Inflated fit	757	683	676	2116
Observation	733	670	701	2104

from certain walk-in centres across the North East. It was after this date that NDUC started to see a growing excess above call forecasts generated from the *CPD* model. Not only did this put extra pressure on NDUC's system at a time when they expect call volume to be low, but it also occurred over the period where most GPs take their holidays (also related to the low demand as well as school holidays). By 9th July NDUC were dealing with an unprecedented call demand on weekdays, and weekends were behaving like the busiest days around the Christmas holidays. Indeed, this activity was occurring all across England, with the HPA's Weekly National Influenza Report published on 16th July 2009 stating that GP consultation rates for influenza and influenza-like illness in England were higher than those observed at the peak of the previous winter, for the week ending on the 12th July. The HPA also had weekly suggestions of how many swine flu cases there were in England, which are archived on their website [?]. Estimates were based on latest weekly consultation rates for flu-like illness, taking into consideration positive swine flu cases (through HPA sampling, as swabbing was no longer the norm), and assumptions related to the volume of patients attending GP appointments. These estimates had large prediction intervals attached to them.

On Thursday 23rd July the National Pandemic Flu Service went live across England, and patients now had two alternative routes for getting antivirals - through a website or dedicated phoneline, and call volume at NDUC started to return back to normal. This unusual activity in July 2009 is clearly visible in figures 3.12 and 3.13 in chapter 3.

5.5 October 2009

In October 2009 call volume began to increase again, and at this time we introduced the method of moving averages to forecast excess call volume.

Let p_{CPD} be the *CPD* predictions for calls per day in October 2009 generated from equation 3.21, and o_{CPD} be the observed calls per day over the same time period. Consider the excess denoted by the letter x , $x = o_{CPD} - p_{CPD}$. Now consider x with respect to the original forecast p_{CPD} , and define the increase ratio

$$I = \frac{x}{p_{CPD}}.$$

5.5.1 Seven-day Moving Averages

Consider the excess x as a time series X_t . Moving averages are smoothing tools designed to eliminate an undesirable component from the series [17]. Next assume that x is a series of a

trend and irregular component. If the trend is smooth, then the values of the series around date t must contain information about the value of the trend at instant t , and it must be possible to use an average of these values as an estimate.

A moving average of coefficients, $\{x_i\}$, is defined as [17, pp 25]:

$$M(X_t) = \sum_{i \in \text{past}}^{\text{future}} x_i X_{t+i} \quad (5.1)$$

The value at instant t is replaced by a weighted average of *past*, past values, the current value, and *future*, future values of the series.

The function `filter` in R allows us to apply a linear filter to a univariate time series. Consider the dataset `excessdata` which runs from 14/09/2009 - 01/11/2009 (inclusive). The R output of `>summary(excessdata)` is shown below. Here $p_{CPD} = \text{fit}$, $o_{CPD} = \text{obs}$, and $x = \text{excess}$. We include some other dummy variables to indicate weekends, holidays, and whether the NPFS was live.

```
> summary(excessdata)
      day      date      dow      fit      obs
Min.   : 1   Min.   :09/14/09   Mon:7   Min.   :110   Min.   :104.0
1st Qu.:13   1st Qu.:09/26/09   Tue:7   1st Qu.:120   1st Qu.:138.2
Median :25   Median :10/08/09   Wed:7   Median :134   Median :162.0
Mean   :25   Mean   :10/08/09   Thu:7   Mean   :247   Mean   :286.3
3rd Qu.:37   3rd Qu.:10/20/09   Fri:7   3rd Qu.:463   3rd Qu.:540.2
Max.   :49   Max.   :11/01/09   Sat:7   Max.   :610   Max.   :769.0
                        Sun:7                        NA's    : 7.0

      excess      h1n1      fluline      weekend
Min.   : -14.00   Min.   : 0   Flu Line:49   Length:49
1st Qu.: 17.25   1st Qu.: 0                Class :character
Median : 30.00   Median : 0                Mode  :character
Mean   : 41.05   Mean   : 0
3rd Qu.: 46.75   3rd Qu.: 0
Max.   :164.00   Max.   : 0
NA's    : 7.00

      pc      pred_excess      ma
Min.   : -6.30   Min.   : 3.00   Min.   : 6.929
1st Qu.:11.10   1st Qu.:21.00   1st Qu.:14.086
Median :20.60   Median :27.00   Median :18.986
Mean   :19.30   Mean   :45.59   Mean   :19.448
3rd Qu.:26.10   3rd Qu.:57.00   3rd Qu.:27.279
Max.   :45.00   Max.   :156.00   Max.   :30.243
                        NA's    : 6.000
```

Now we want to predict x , and assess the effects of *day* (the number of the day of the series with 14/09/2009 = 1), and *dow*, and a possible interaction between the two. Limiting to this date range, we wanted to assess the imminent swine flu impact for the next week. Consider the following linear models defined in equations 5.2 and 5.3

$$x = \beta_0 + \beta_1 \text{day} + \beta_2 \text{dow} + \epsilon \quad (5.2)$$

$$x = \beta_0 + \beta_1 \text{day} + \beta_2 \text{dow} + \beta_3 \text{day} : \text{dow} + \epsilon \quad (5.3)$$

The Analysis of Variance table below in the R Output showed that, at this stage, and using these exclusions, equation 5.2 is the favourable model, which is supported by the residual plot in

figure 5.2. In this particular residual plot we have coloured the days ensuring there is no pattern, and the interaction term can be omitted. The only alarm bell is that Sundays (black points) seem to have a large variation, so we need to recheck the model as new Sunday observations arrive.

```
> anova(predexcess1,predexcess2)
Analysis of Variance Table

Model 1: excess ~ day + dow
Model 2: excess ~ day * dow
   Res.Df  RSS Df Sum of Sq    F Pr(>F)
1      34 25485
2      28 18112  6   7372.6 1.8996 0.1161
```

After we have established a way to predict x , we do so for the next seven days. Figure 5.3 illustrates this excess activity by plotting the observed excess, in green, for past days, and the predicted excess, in blue, for future days. The predictions are generated from equation 5.2. At this stage you can see that the excess was increasing, and we predicted about 50 excess calls on weekdays, and over 100 excess calls for weekends on this particular week.

Finally, we take the excesses, and calculate the percentage increase from p_{CPD} , or `fit`. This is the field `pc` in the summary of `excessdata` in the earlier R output. Finally, we define the seven-day moving average (`ma`) with the R code:

```
> excessdata$ma<-filter(excessdata$pc,rep(1,7)/7)
```

Figure 5.4 plots these moving averages, with observed in green, and predicted in blue, as before. This is what gives NDUC an idea of the imminent excess being predicted over a week. At this particular time we were predicting weekly excesses of about 35%.

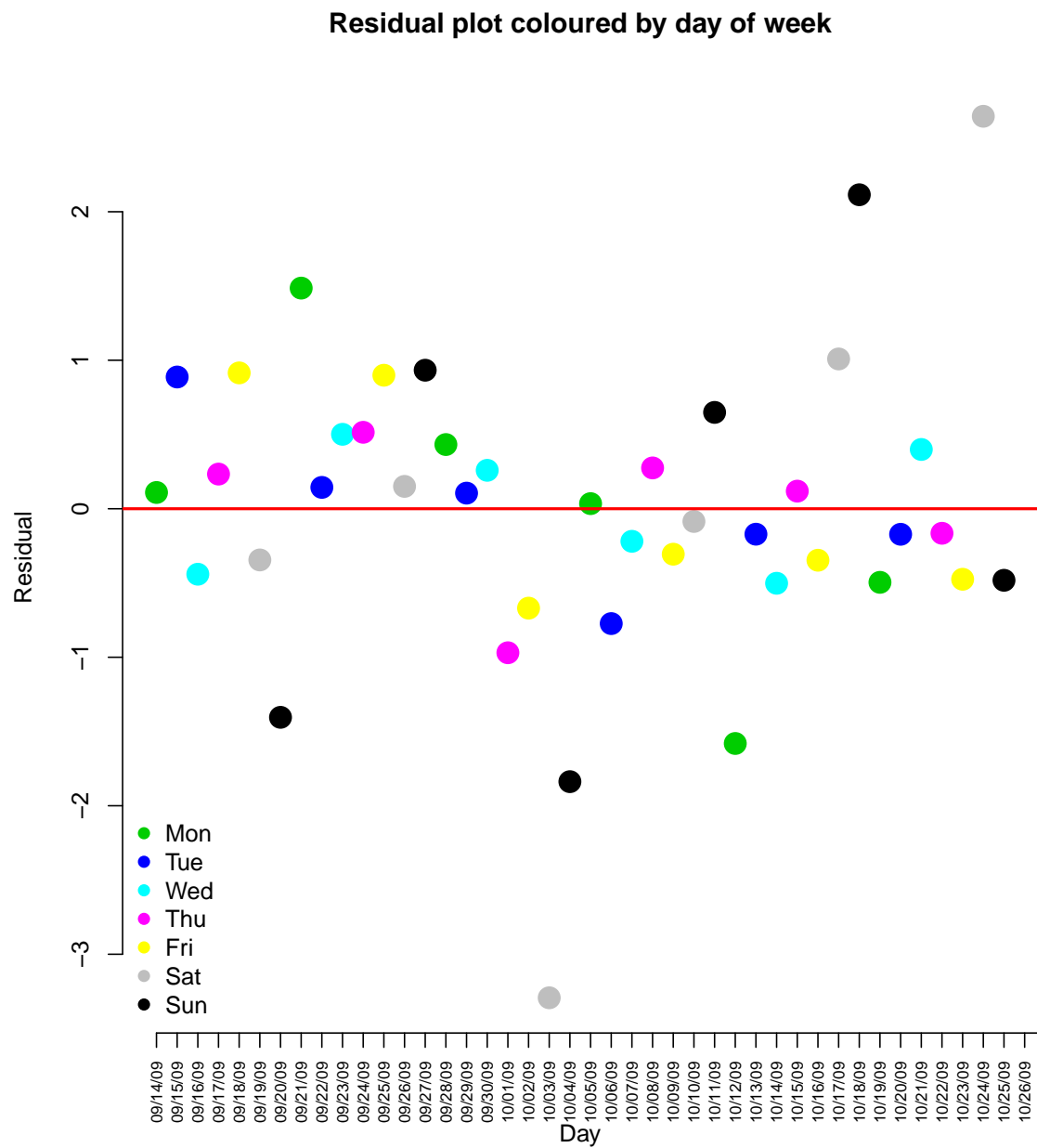


Figure 5.2: Standardised residual plot of equation 5.2

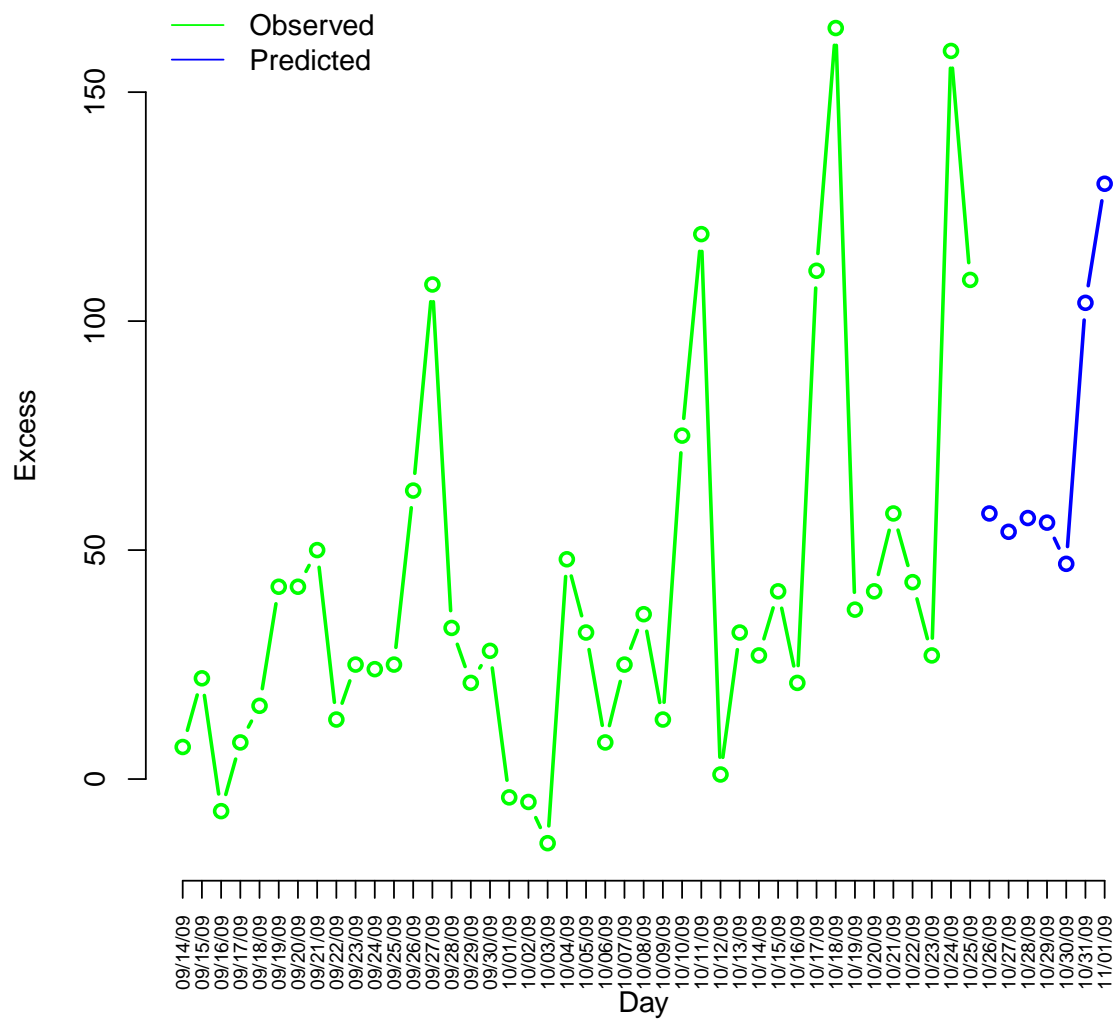


Figure 5.3: Predicting excess from equation 5.2

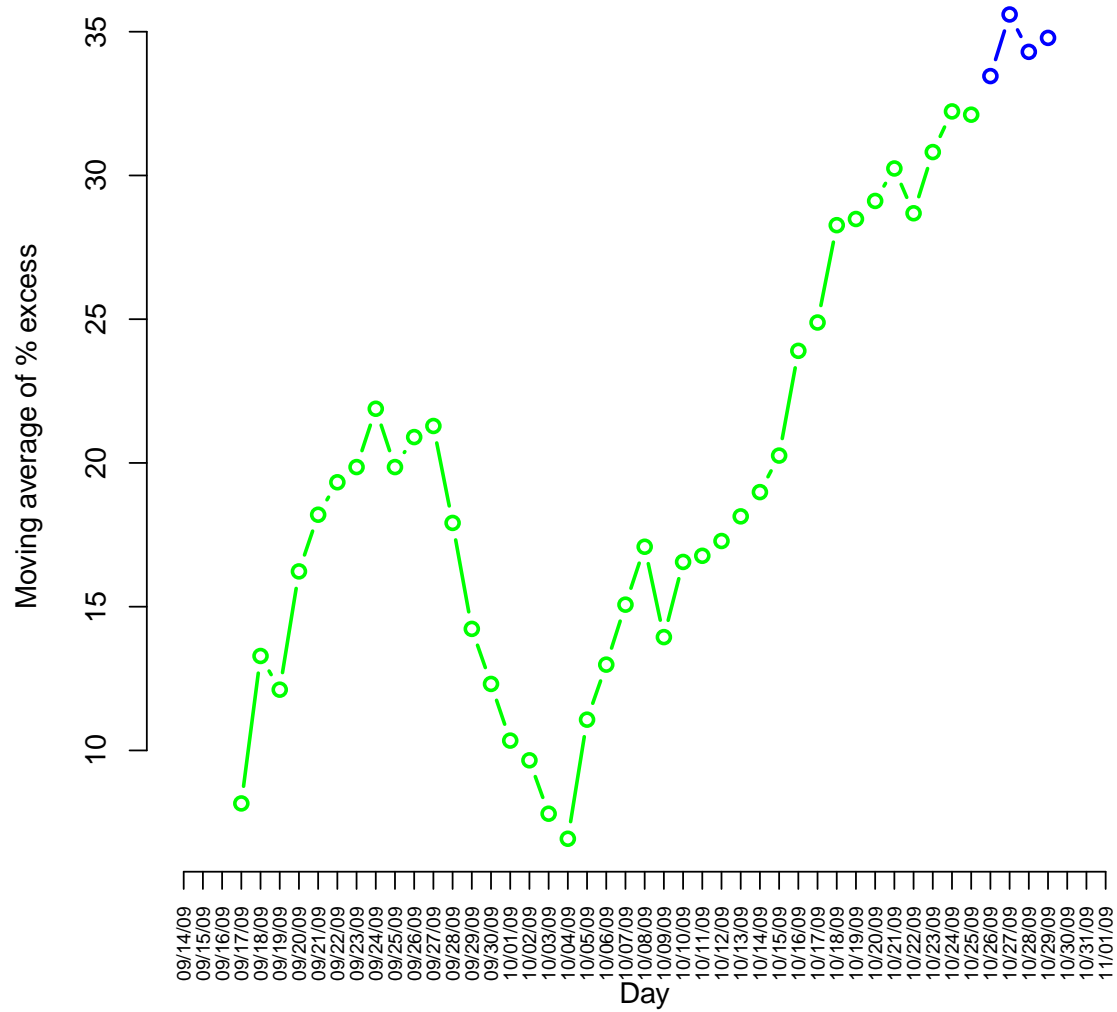


Figure 5.4: Applying a seven-day moving average

5.5.2 North East Strategic Health Authority

On Monday 19th October most of the out-of-hours providers, and several PCT representatives from the North East group met to discuss the swine flu pandemic at the North East SHA. Six outcomes were agreed at this meeting and our work contributed to two of them:

1. Out-of-hours providers should define their escalation policy, using the North East Escalation Plan (NEEP) 1-6 framework
2. Provide timely out-of-hours activity data. This would help as an early warning trigger for the rest of the healthcare system.

5.5.3 North East Escalation Plan

The idea behind NEEP was to implement a Pandemic Flu Escalation Framework, with clear triggers and actions, so the organisation knew what was expected of them at every level of pressure. The structure of the report a simple spreadsheet as illustrated in figure 5.5, and we had to complete this form clearly and logically. The structure of the plan is that it defines the organisation to be at one of six potential levels:

- Level 1: Normal (white)
- Level 2: Concern (green)
- Level 3: Pressure (amber)
- Level 4: Severe Pressure (red)
- Level 5: Critical (purple)
- Level 6: Potential Service Failure (black)

For each of the six levels you had to define six variables:

1. Trigger: What needs to have happened, or is expected to happen to move to this level?
2. Action: What will be done to mitigate the raised level of pressure as a result of this level?
3. Communication: What will be communicated internally and externally?
4. Command/control: Who has the authority to move to this level?
5. Impact: Expected impact of actions
6. Implications: Will actions affect other organisations?

Our main input was through the first variable, Trigger. Given that we had an accurate *CPD* model, what sort of levels above *CPD* forecasts will result in a move to NEEP levels 2, 3, 4, 5, and 6? This was particularly difficult to assess, there wasn't an obvious way to approximate the number of excess calls as external factors, like time of year, and holidays, will affect it. The document also required the increase to be written as a percentage. Then we needed to decide what was the appropriate level to calculate this trigger? Daily? Weekly? Monthly?

Initially there was favouritism towards having a daily outlook, but as call volumes can be low on weekdays a daily percentage increase or decrease trigger is not reliable. Monthly was not viable as by the time you work out whether your activity was normal or not, the pressure period is likely to have subsided. We proposed the best way to look at this would be similar to the *CPD* model, but this time apply that logic for forecasting calls per week from Monday

Trigger	Level	Action	Communication	Command/control	Impact	Implications?
What needs to have happened? or be about to happen?		What will be done to mitigate the raised level of pressure as a result of moving to this level? Who by? When? Where?	Internal and external communications Who by? When?	Who has the authority and responsibility to trigger?	From actions	Any implications of actions on other organisations?
	1 normal					
	2 concern					
	3 pressure					
	4 severe pressure					
	5 critical					
	6 potential service failure					

Figure 5.5: NEEP Framework

- Sunday *CPW*. Furthermore, why not take the previous seven-days forecasts from the *CPD* model, aggregate them, and compare it to what's actually happening? That way you don't have to wait until Monday morning to see how the previous week behaved. This approach was also more convenient during a period when there were severe time pressures, and the organisation was already familiar with the *CPD* model. This approach was adopted, and appropriate triggers were chosen to complete NDUC's NEEP policy, which is still in place today. The triggers for the six NEEP levels are detailed below, and are percentage increase on calls per week, depending what season it is.

Level	Trigger
1	No trigger
2	17% – 23%
3	23% – 41%
4	41% – 53%
5	53% – 100%
6	> 100%

5.5.4 SHA Situation Reports

SHA Situation Reports, or SitReps, were created to provide timely out-of-hours activity data. Between 28/10/2009 and 28/02/2010 the activity return form illustrated in figure 5.6 was completed and sent to the SHA contact, by 11:00 every morning, for the previous shift, firstly, seven days a week, and then five days a week. Moreover, we had to complete two of these; one for North of Tyne (Northumberland, North Tyneside, and Newcastle) and South Tyneside.

There were several issues with this form, the first being the frequency. Since the report had to be completed seven days a week, and couldn't be done remotely, this responsibility fell to the Operations manager who was on call at weekends, so a user friendly process had to be created. The next issue was that we had to split North of Tyne and South Tyneside. In Q1 in figure 5.6, we had to input the number of calls received, that is, calls that are going straight to a call handler. In practice, this will always be higher than the number of out-of-hours cases created for GPs to triage. Furthermore, with our telephone system at the time, we could not distinguish where the calls were coming from; this classification is only recorded once a GP out-of-hours case has been created. Therefore, we could only report on total calls coming into our system, from all areas.

Then there was the forecasting; the *CPD* model forecast the number of genuine out-of-hours cases for the whole organisation, and wasn't split for area. Forecasts for home visits and centre

	A	B	C	D	E	F	G	H	I	J
1										
2			NHS North East: Out-of-hours Provider Activity Return							
3										
4				OOH Service	<input type="text"/>					
5										
6				Reporting period - 24 hours prior to 08:00am on:	<input type="text"/>					
7										
8										
9			Q1	Total number of calls received in period	<input type="text"/>					
10				% variance from organisation forecast	<input type="text"/>					
11										
12			Q2	Total number of calls triaged in period	<input type="text"/>					
13				% variance from organisation forecast	<input type="text"/>					
14										
15										
16			Q3	Total number of home visits undertaken in period	<input type="text"/>					
17				% variance from organisation forecast	<input type="text"/>					
18										
19										
20			Q4	Total number of centre visit consultations in period	<input type="text"/>					
21				% variance from organisation forecast	<input type="text"/>					
22										
23										
24			Q5	Optional comments section	<input type="text"/>					
25										
26										
27										

Figure 5.6: SHA SitRep template

visits had large error bounds, and again were not split by area. We didn't split forecasting down to this level as it was organisationally unnecessary, and could result in misleading statistics. Consider the following example to illustrate my point; home visits on Tuesdays in October in South Tyneside. We expect there will be 3 home visit consultations on this particular day, however, it would not be unheard of to see 0 or 1, or 5 or 6 consultations on a day like this either. Reporting with the return in figure 5.6, we were occasionally saying that on certain shifts there were 100% increases or decreases in activity levels, which is of course, misleading to the reader.

Not only this, but the report was completed daily, and this was supposed to guide you to choosing your NEEP level. Returning to home visits on Tuesdays and Wednesdays in October, in South Tyneside, we would forecast 3 on each day. Say we observed 1 on Tuesday and 5 on Wednesday we would report a 66% decrease followed by a 66% increase, which would take our organisation from NEEP Level 1 to NEEP Level 5 in one day, when activity is just varying as we would normally expect.

These issues were raised to our commissioners and it was clear there was some confusion over the NEEP framework. For example we would report a daily decrease, but have a weekly increase over the latest seven days warranting a move to NEEP level 2, and this decision would be questioned. Then it was a case of reporting a daily increase but the week had behaved normally, and we would be questioned on our decision to remain at NEEP level 1. We spent some time explaining this method to the commissioners, and how we were assessing activity differently, and they were happy to let us continue to do it on a seven-day basis. This open communication built on already existing relationships with the commissioners, and we became recognised as a reliable data source. Even though we identified the flaws in reporting like this, and were mostly supported in our comments, the daily SHA SitRep remained in place for just over four months, and the time invested in this could probably have been put to better use elsewhere.

Chapter 6

Performance Monitoring

Central to operations at NDUC, it is necessary to monitor the activity, average handling times, and case outcomes of all the staff dealing with calls at the call centre. To meet the organisation's KPIs, it is necessary to make some assumptions, or expectations, around average handling time (AHT), and how many calls they are expected to take in a certain time period, given appropriate demand. This is more complicated to implement among the GPs, given the complexity of the operational setup, and patient case mix.

This chapter discusses the resources available to managers for monitoring their staff on shift before the KTP project, and the tools put in place to improve their management information, for both call handlers and GPs. It then goes on to discuss the predictive methods implemented when trying to model individual GP's behaviour on shift.

6.1 Call handlers

After the recorded message, the first port of call for the patient is to speak with a call handler. Here the call handler spends around 3 minutes taking personal details of the patient, a summary of symptoms, and assigns an urgent or routine priority. Before making it through to a call handler, the patient takes another journey; they listen to a recorded message, which is 19 seconds long. This explains the service they are about to receive, and that a call handler will pick up the call shortly following the message. The patient will then be put on-hold, waiting for a call handler to become available. After holding, they then speak to the call handler. During this journey the patient has two slots to leave the system before reaching a call handler:

1. patient hangs up during the recorded message; call handler had no chance of picking up
2. patient hangs up while on-hold after the recorded message; this is classed as an abandoned call
3. patient gets through to a call handler after being on-hold; this is classed as an answered call

Every month NDUC report to their commissioners on two KPIs concerning abandoned calls and on-hold times:

1. Following the end of the recorded telephone message, the abandoned calls rate should be 5% or less
2. 95% of answered calls should be answered within 60 seconds or less after the end of the recorded message.

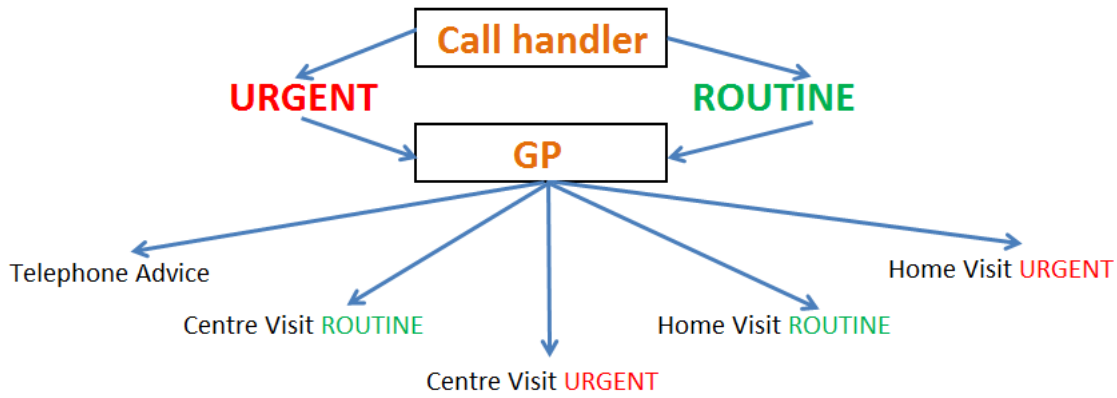


Figure 6.1: Workflow of call handlers and GPs

We also note here that call handlers have a direct impact on the GP KPIs. This is because call handlers classify calls as urgent and routine, and these cases carry different KPIs (as discussed in chapter 2 section 2.4.1 where urgent calls need to start definitive clinical assessment within 20 minutes, and routine within 60 minutes). The higher the rate of urgent calls generated by the call handlers, the greater the pressure on the GP triage workload. The different stages of the patient path from call handler to GP is illustrated in figure 6.1.

For efficiencies of staffing we also need to have a realistic target for AHT, and communicate that to staff. At the start of this research there was little information about call handler behaviour, with none on a shift or individual level. Throughout this project time was spent understanding the program MyCalls; this allowed extraction of raw data from which we could generate summaries around shift performance and call handlers' AHT. This information did not progress past basic summary level, but the visibility of performance, shift and server wise, still added value to NDUC as an organisation. Further work could develop the data extracted for call handlers, and try to understand queuing from the front end of the system; the telephone switch.

6.2 GPs

The research described in monitoring GPs, particularly around their triage times was collaborative work with David Wooff at Durham University.

GPs are very expensive. The difference in cost between staffing three GPs on weekday evenings versus four amounts to around £70,000 p.a. Therefore it is imperative we have accurate forecasts to produce demand-driven staffing levels. These, however, are not enough. Experience tells us that GPs are not interchangeable, and they will all have an individual productivity level, dependent on the individual GP and external factors, such as the size of the queue. Varying GP productivity will affect the response time to patients; if you have a group of GPs with low productivity all working on one shift, patients will have longer waits before contact than if you have a group of GPs with high productivity.

First we need to split the GP resource into two categories; triage, and face-to-face. Triage is a telephone conversation between a GP and patient, where they assess the patient's needs. Face-to-face contact is where a GP assesses a patient either at one of the seven UCCs or a patient's home. Every OOH case will be triaged, but not every OOH case will require a face-to-face consultation. Therefore in what follows we restrict GP productivity to considering only triage activity, with obvious extension to other activity.

Over any shift there will be different types of GP shift, where GPs have to undertake a combination of roles. There are shifts based at headquarters where GPs will triage calls and perform face-to-face home visits. There are shifts based at our UCCs where GPs will consult with patients through centre visits and referrals from A&E departments, as well as triaging calls remotely when there is enough demand. There are shifts where GPs triage remotely, from their home, for a short period of time. Depending on the GPs surroundings, this may have an impact on their productivity.

Furthermore, it is intuitive to think that the number of people in the system at any time, the size of the triage queue, will also affect GP productivity. If the GP is triaging at headquarters and there are 20 patients waiting for a call back with a steady arrival rate, you may expect that GP's productivity to be higher than if there were only one or two people sitting in the queue. That is, if there is no queue, the GP can afford to spend longer triaging a call, and tends to do so.

For every case triaged, the GP decides the outcome (recall figure 6.1): telephone advice, centre visit, or home visit. Home visits are the most expensive outcome, followed by centre visits, and then telephone advice. Naturally, you would expect different GPs to have different rates of telephone advice versus home visits, however, there will also be GPs who naturally lean towards deciding on a home visit rather than telephone advice. High home visit rates also put pressure on the system further down the line. Not only this, you could expect final outcome to affect the length the GP spends triaging the call. We would typically expect calls which turn out to result in visits to be shorter in length.

We have briefly discussed a few issues as to why monitoring and modelling this productivity is difficult. Given the nature of the setting, each out-of-hours case represents a patient, not a number, and there is a duty of care to all patients who come through the system. Acknowledging this fact, it is still essential to monitor GP behaviour, identify outliers for retraining purposes, as well as trying to model their productivity for future planning.

The next two subsections show techniques developed and implemented during the KTP project, which deal with the issues above.

6.2.1 Monitoring Productivity over Shifts

Given the setup of GP operations at NDUC, and that often GPs are working remotely from headquarters, it is necessary to monitor when they are being active, and when they are not. We created a timeline of the shift, with each GP's activity illustrated individually, colour coding it for different types of patient contact. Here you could identify whether certain GPs have gaps in their activity. Then we had to assess the reasons behind their activity pattern. Maybe there are gaps in their activity because there were no patients waiting to be triaged. Therefore we created functions in R to generate the GP activity for each shift, and to visualise the triage queue over the same period. To do this, define five types of shifts on any weekday:

- *Triage/Home Visit* shift (Tri/HV). Here the GP is based at headquarters for the evening part of a weekday shift. Their shift will start at time t_1 and finish at time t_2 . The GP is responsible for triaging calls, as well as attending home visit consultations.
- *Urgent Care Centre* shift (UCC). There will be seven UCC shifts in any weekday evening shift. Here, one GP is based at one of the UCCs NDUC operates. These GPs are responsible for centre visit consultations, referrals to them from an A&E department, and triaging calls from the UCC as and when required.
- *Overnight* shift (o/n). Here, the GP is based at headquarters for triaging calls and attending home visit consultations.
- *Nurse* shift. Nurses work at headquarters during weekday evenings, dealing with dental out-of-hours calls as well as routine GP out-of-hours calls.

- *Home triage* shift (H/triage). Here, the GP is based at their home, and triages calls remotely. They perform no face-to-face consultations.

We order the shifts as described above for convenience. It should also be noted that different shifts can be done by the same GP. For example a GP may be working at a UCC in the evening, and then return to headquarters to start an overnight shift.

6.2.2 The GP Activity Plot

Figures 6.2 and 6.3 show the GP activity plot and key, respectively, used at NDUC for a shift on Tuesday 25th May 2010. There are seventeen rows in the picture in total. Horizontal lines represent some sort of activity with a patient, and end points mark the end of this activity, as well as giving an indication on the case type. Triage activity is denoted by the light blue horizontal line, and routine home visits and urgent home visits are denoted by the green and red horizontal lines respectively. Details of the meanings of the colours and end points are detailed in figure 6.3.

You can see the picture looks quite complex, but if you split the GP resource into Tri/HV, UCC and (o/n) it becomes less so. You should also note there can double shifts over a weekday evening out-of-hours session. The GP working at UCC 7 finished their work, and then travelled back to headquarters to start an overnight shift.

Consider figure 6.2 starting from the top and working down. There were four GPs based at headquarters responsible for triaging and home visits. You can see that the first GP was scheduled to be on shift between 19:30 and 01:00. They started their shift with some telephone triage (5 calls in total), then went out to consult two routine home visit consultations, and then returned to headquarters to triage a further three calls.

Compare this with the second Tri/HV GP who was scheduled to be on shift between 19:00 and 23:00. This GP triaged pretty much solidly throughout their shift, handling about 3 calls an hour between 19:00 - 21:00, and then speeding up to 6 or 7 calls an hour between 21:00 - 23:00.

Underneath the Tri/HV GPs you have the seven UCC GPs. Consider now UCC 2 and UCC 4. The GP at UCC 2 started the shift by triaging eight calls. He moved on to a routine centre visit consultation, before triaging another call. There was then a gap in activity for 30 minutes before he finished his shift by alternating between triage and routine centre visit consultations. The GP at UCC 4 recorded one centre visit routine consultation, over the whole shift, which lasted for nearly four hours. It is highly unlikely that this is accurate; however, face-to-face consultation times are entered by the user, which results in inputting errors from time to time. In reality, we don't know what the GP at UCC 4 was doing, but this highlights that something strange was going on, perhaps to do with technical issues.

Next we move onto the overnight shifts, and you can see four GPs mostly triaging calls, and a few going out to home visit consultations. There are regular gaps in activity.

To finish we have the nurse on shift, who handles a mixture of dental and GP calls, followed by a few consultations coming up at one of our walk-in centres (Jarrow) that closes at 20:00, and a GP triaging two calls at around 23:15 from their home.

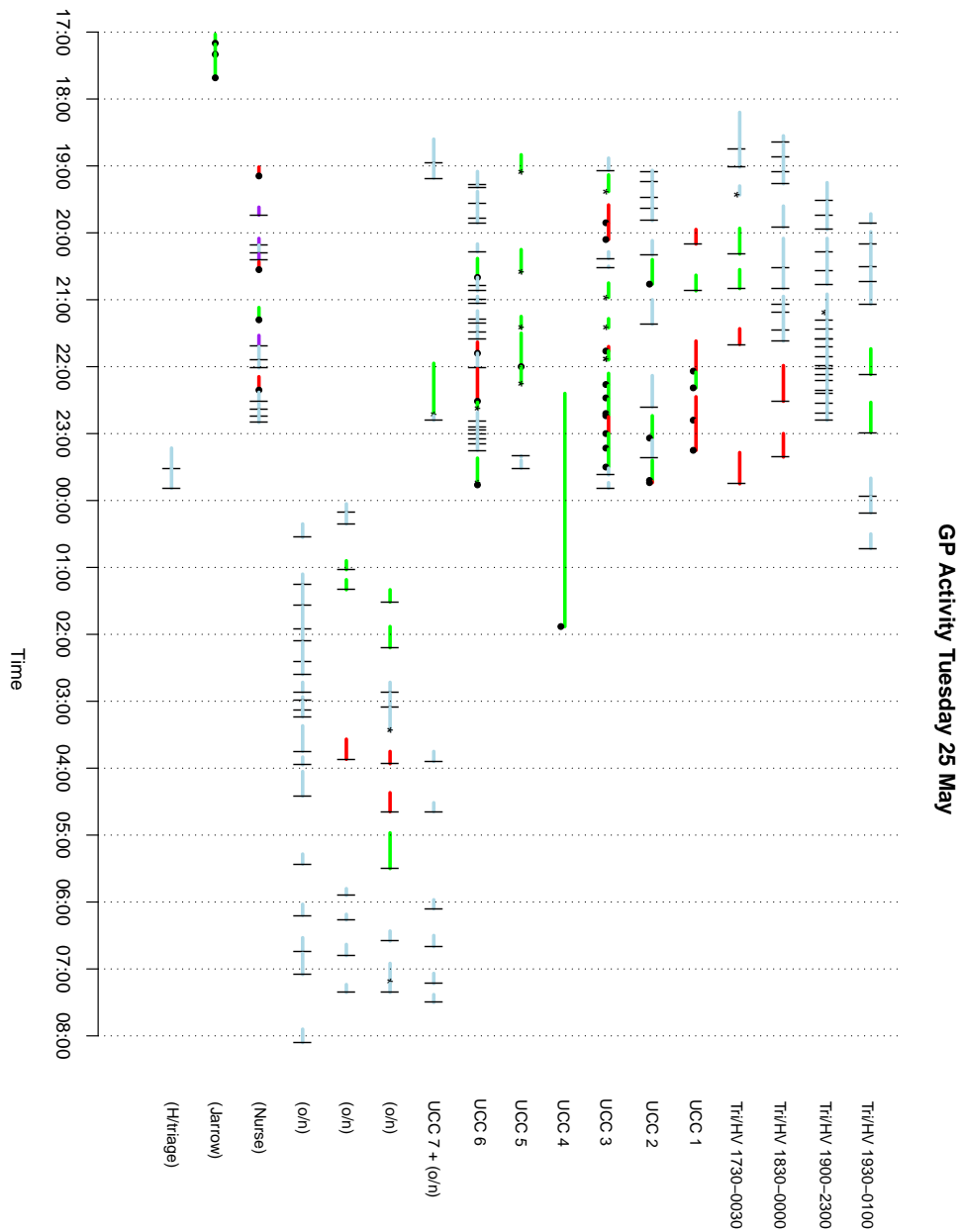


Figure 6.2: GP activity plot

GP Activity Graph guide

The following graphs can be produced on a shift by shift basis.
The key to read them is set out below. Remember:

1. Lines represent contact with patient
2. End points represent outcomes
3. Nurses are based at N/House for routine triage/dental calls – their timeline shows **the outcomes of their triage** (coded below).

	Telephone triage	• Tri/HV: GP triage and home visit shift
	Other triage (incl NEAS support)	• STP: South Tyneside UCC
	Routine face-to-face Home Visit	• NTP: North Tyneside UCC
	Urgent face-to-face Home Visit	• NGH: Newcastle General Hospital UCC
	Routine face-to-face Centre Visit	• HEX: Hexham UCC (note A&E covered on evenings 1900-0800)
	Urgent face-to-face Centre Visit	• BER: Berwick UCC
	A&E Referral	• ASH: Ashington UCC
	Dental call	• ALN: Alnwick UCC
		• o/n: Overnight shift
		• H/triage: Home triage available
		• H/triage sched: Home triage scheduled
		• Nurse: Nurse (also do dental calls).

Figure 6.3: GP activity plot key

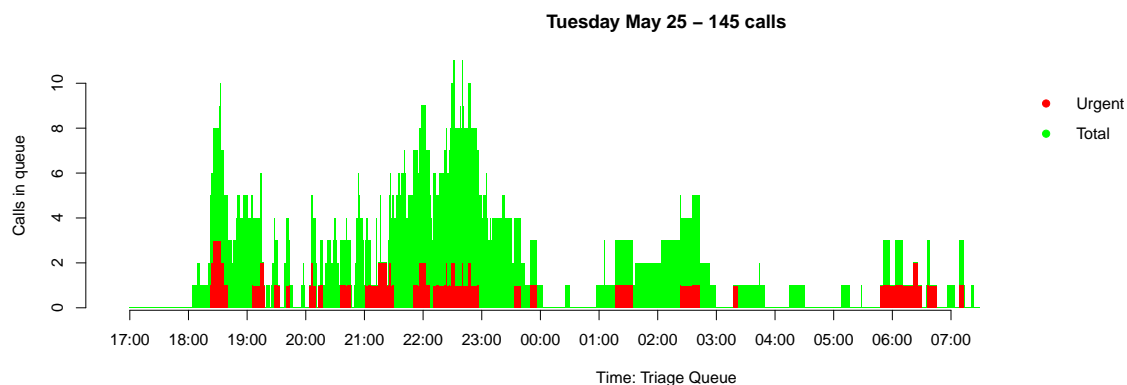


Figure 6.4: Triage queue plot

6.2.3 The Triage Queue Picture

Consider now, figure 6.4, which shows how many routine and urgent cases were waiting to be triaged on Tuesday 25th May. As with the GP activity plot, the triage queue picture is broken down to minute level. The red bars show how many urgent cases are in the queue at any time, and the routine calls (in green) are stacked above these to give the total number of cases waiting to be triaged at any minute of the shift.

Recall the second Tri/HV GP from figure 6.2, and their varying work rate between 19:00-21:00 and 21:30-00:00. If you look at figure 6.4, the triage queue was most pressurised between 21:30 - 00:00, which could have driven this GP to speed up their triage. On the overnight part of the shift the queue is mostly empty, so it's not unusual to see gaps in activity during this time.

The GP activity plot and triage queue pictures have been produced daily at NDUC since August 2009, and still make up a valuable piece of information when assessing shift performance and understanding GP behaviour. As yet, we have not formalised a link between GP activity and the triage queue, but future work could address this issue.

6.2.4 Monitoring Triage Lengths

NDUC have two conflicting issues to balance when monitoring GP productivity, or triage length.

1. To reach optimum staffing we need to know arrival rates and GP productivity. Ideally we want GPs to work as quickly as possible.
2. Given the nature of the calls, NDUC cannot say “all GPs must deal with 5 calls per hour”. Different case mixes will take different lengths of time, and there is duty of care towards patients. On a lengthy case it is unacceptable to reach the ‘time limit’ of 12 minutes and cut the case short. We know GPs will take longer or shorter on triage depending on their conditions, ability and personal attributes. Particular GPs, like particular members of any workforce, will be lazier than others, and it is important to be able to identify the correct reasons for the differences in speed, and monitor accordingly.

The investigation began by exploring a cohort of GPs compared to each other, rather than setting KPIs. For example if GP X and GP Y always work the same shifts, with the same case mix, with the same experience, and GP X is always slower than GP Y, we need to know why. Therefore this research will:

- Ensure there is an appropriate skill mix on each shift

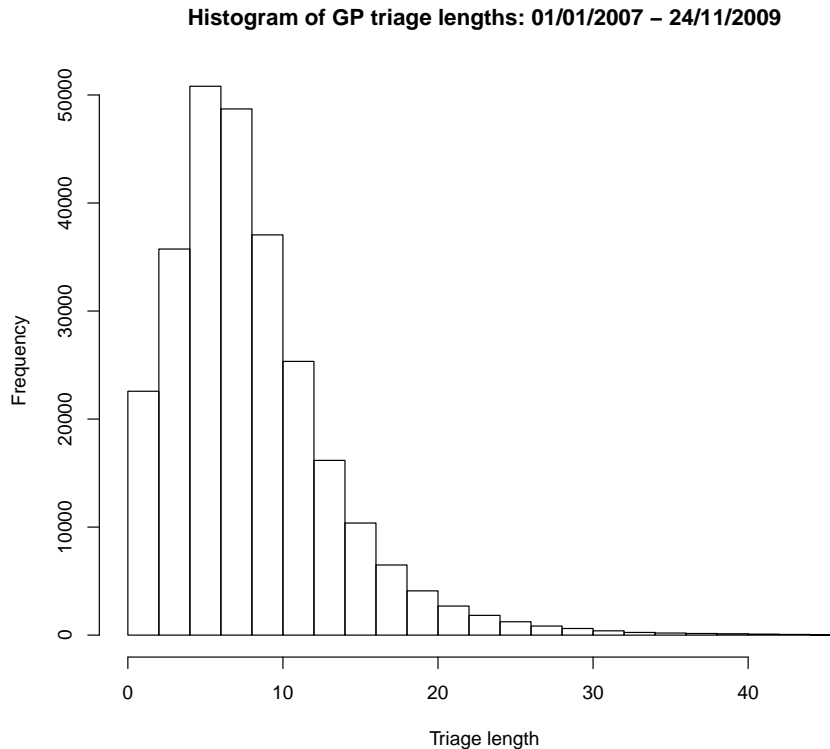


Figure 6.5: Histogram of GP triage lengths

- Highlight outliers for call reviews and possible retraining

In the first instance, all out-of-hours cases were collected from January 2007 until 24th November 2009. From this, all telephone consultations, or triages, by each clinician were recorded (over 250,000 such consultations). We cleaned the data by reducing to certain cases in Adastra:

- Triage has a consultation type of ‘TELEPHONE ADVICE’ or ‘Streamed - Advice’. This ensures such non-OOH triages, like North East Ambulance Service Clinician Support, are removed.
- Triage has a final case type of ‘TELEPHONE ADVICE’, a derivative of ‘Centre Visit’, or ‘Home Visit’, and ‘Patient Rejected Advice’. This ensures non-out-of-hours cases, like ‘Telephone Answering’, are removed.
- Triage GPs are fully qualified. Some calls are handled by Registrars, who are not fully qualified, and may triage under supervision.
- GPs have triaged at least n cases, to ensure estimates are reasonably robust. In the method that follows, $n = 20$

Triage lengths are rounded to the nearest minute, and we use exploratory data analysis to establish the distribution we will use for GP triage lengths.

Let y denote GP triage lengths, in minutes. Figure 6.5 shows the distribution of y , which is skewed as might be expected. When you have skewed observations, one possibility is to transform to Normality before-hand. In this case, taking logs works well. However, a disadvantage is that models which fit to a transformed response variable cannot easily be back-transformed in

order to give results on the original scale [15]. A superior alternative in this case, as described below, is to use a fitting technique which can take explicit account of the shape of the response variable.

Next we needed to know what would affect y . All GPs triage two types of calls, routine and urgent. Intuition would tell us that urgent triages will take longer than routine triages, as they are more complicated. There were also clear time of day and time of year effects. For time of year, triage lengths are slightly longer between April-November, than December-March. We also know, however, that the service is much busier in the winter months, and a GP faced with a busy shift and large queue will work faster than a GP faced with a quiet shift and a short queue - that is, they can afford to take more time in Summer than Winter due to demand on the service. Similarly, for time of day, triage lengths are highest on overnight periods (00:00:00 - 07:59:59), and shortest in the morning times. Again we also know that the morning time of day (08:00:00 - 11:59:58) only occurs on Saturdays and Sundays, the busiest time of the week, and overnights tend to be the quietest time.

As well as this, feedback from the Head of Clinical Services, who still works out-of-hours GP triage shifts, suggested that different days and times throughout the week generate different caseloads. Saturday mornings will generate a large number of routine cases, filled with simple ailments and repeat prescription requests. Overnights will generate a small number of more complicated cases. This knowledge backs up the findings in the model. A sensible way to deal with these effects was to leave them out of the modelling, and subset the dataset used in modelling to account for them. For example, when considering in April how GPs have been behaving recently, take data from January-March, and limit to particular times of day, like Saturday mornings in January - March. Then model GP behaviour from this dataset.

The GP also has a choice about the outcome, or case type, of the triage. This will either be telephone advice, centre visit, or home visit. During the first minutes of triage it will be apparent if a face-to-face contact is appropriate. If that is the case one will be arranged and the GP can assure the patient that another GP will follow up on their case. If that is not the case the GP will have to reassure the patient with advice, which will take longer than saying "A receptionist will call you back shortly to arrange a centre visit", or "A GP will call with you at your house within two hours". The different mix of outcomes will also reflect different days and times. The UCCs are closed on the overnight time periods, and so the only possible outcomes are telephone advice or home visit. This theory was tested and indeed, outcome was an important correlate.

For the sake of simplicity we choose to fit a basic model which adjusts for whether or not the case was urgent as well as the outcome. Then we apply different subsets of the dataset to run through this model. For example run the model for the clean dataset for triages on *Saturday mornings* (Saturday 08:00 - 11:59 excluding the Saturday before a Bank Holiday), *weekday evenings* (Monday-Friday 18:30 - 23:59), *overnights* (Monday-Sunday 00:00 - 07:59), and *bank holiday weekends*.

A generalised linear model, with gamma errors and a log link works in practice. This gives the model of the form

$$\hat{y}_i = \exp(\alpha + \sum_i \beta_i x_i),$$

where the x_i are dummy variables with respect to urgency and outcome. This, in turn, gives us the model:

$$\hat{y}_i = e^\alpha \prod_i e^{\beta_i x_i}.$$

The proportionate effect of x_i on the intercept is a simple interpretation of the relevance of x_i .

Consider all GPs triages in the second quarter of 2010 (April - June inclusive). The R Output below shows the generalised linear model as described above, assessing the importance of urgency (`urg`) and outcome (`casetypeb`) on triage length (`trriage`). The baseline is a routine call ending as telephone advice, and as we expected, face-to-face outcomes are all shorter, and urgent calls are longer. The average triage time is nearly 00:10:30, and the practical implications of these effects are under a minute for outcome, and about 00:01:20 for urgent cases.

Call:

```
glm(formula = triage ~ casetypeb + urg, family = Gamma(link = log),
     data = data)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.7334	-0.4778	-0.1115	0.2428	1.9247

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.33157	0.02424	96.205	< 2e-16 ***
casetypeb[T.Centre Visit ROUTINE]	-0.42619	0.03307	-12.888	< 2e-16 ***
casetypeb[T.Centre Visit URGENT]	-0.21945	0.06672	-3.289	0.00102 **
casetypeb[T.Centre Visit With Transport]	-0.14038	0.08892	-1.579	0.11456
casetypeb[T.Home Visit ROUTINE]	-0.15593	0.03992	-3.906	9.7e-05 ***
casetypeb[T.Home Visit URGENT]	-0.09061	0.06095	-1.487	0.13728
urg[T.Urgent]	0.08116	0.03590	2.261	0.02390 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.3459581)

Null deviance: 746.21 on 1857 degrees of freedom
 Residual deviance: 682.66 on 1851 degrees of freedom
 AIC: 10860

Number of Fisher Scoring iterations: 5

Analysis of Deviance Table

Model: Gamma, link: log

Response: triage

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	P(> Chi)
NULL			1857	746.21	
casetypeb	5	61.733	1852	684.48	< 2e-16 ***
urg	1	1.819	1851	682.66	0.02184 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

ALL

With summer and winter months

Average triage time: 10.29 minutes. Percentage multipliers are:

casetypeb[T.CV ROUTINE]	casetypeb[T.CV URGENT]
-34.70	-19.70
casetypeb[T.CV With Transport]	casetypeb[T.HV ROUTINE]
-13.10	-14.44
casetypeb[T.HV URGENT]	urg[T.Urgent]
-8.66	8.45

Practical impact of the multiplier on the average
is a difference of these times in minutes:

casetypeb[T.CV ROUTINE]	casetypeb[T.CV URGENT]
-0.99	-0.90
casetypeb[T.CV With Transport]	casetypeb[T.HV ROUTINE]
-0.76	-0.80
casetypeb[T.HV URGENT]	urg[T.Urgent]
-0.61	1.31

Our interest is in the activity of individual GPs when triaging, in particular: is the GP faster or slower than average, and how variable are their triage lengths? The focus now moves to the residuals from the model. The model adjusts for factors we know we need to take into account, and the residuals now summarise GP behaviour. We examine this behaviour by looking at boxplots of these residuals. Through the boxplots, we can identify three groups of GP

1. GPs with 75% of their cases below the average triage length i.e. ‘quicker’ GPs.
2. GPs with 75% of their cases above the average triage length i.e. ‘slower’ GPs.
3. GPs performing similarly to the average.

Doing this will ensure very long triage lengths (e.g. due to mental health cases) and very short triage lengths (e.g. due to engaged calls) do not affect the GP rating. It is also important to consider how many cases each GP has triaged. In practice this will be illustrated by the width of the box: the wider the box the more cases.

Figures 6.6 and 6.7 illustrate this boxplot method for residuals for weekday evenings, and Saturday mornings, in the second quarter of 2010, respectively. ‘Quicker’ GPs appear to the left of the graph and ‘slower’ GPs to the right. GPs behaving similarly to the average are omitted from the picture (33 in the weekday case, and 27 on Saturday mornings). Now consider these pictures carefully.

Does it seem that experienced GPs are on the left and inexperienced GPs are on the right? In the weekday case the slow GPs have narrow boxes, indicating they have triaged fewer cases, but quick GPs have a mixture of wide and narrow boxes. Slow Saturday GPs seem to have narrower boxes than fast Saturday GPs. Is there typically the same pattern over different shifts? In both pictures, SGP1D comes out at the extreme left, and no slow GPs are repeated in each picture. Perhaps this is another indication of experience.

This process was piloted at NDUC for the second quarter of 2010 (see figures 6.6 and 6.7) with good results. We found that we weren’t telling the company anything they didn’t know, but provided evidence of what was once anecdotal, to feedback to the GPs. This was used to identify slow GPs so the clinical team could perform call reviews, which influenced training of those GPs.

In practice, it is recommended to re-run this process every three months. This will ensure seasonal effects are kept to a minimum and can be omitted from modelling, and that GPs aren’t classified as ‘slow’ or ‘quick’ for a long time without being able to change it, particularly after addressing training needs.

Weekday evening triages: differences in GP behaviour; 20+ cases per GP

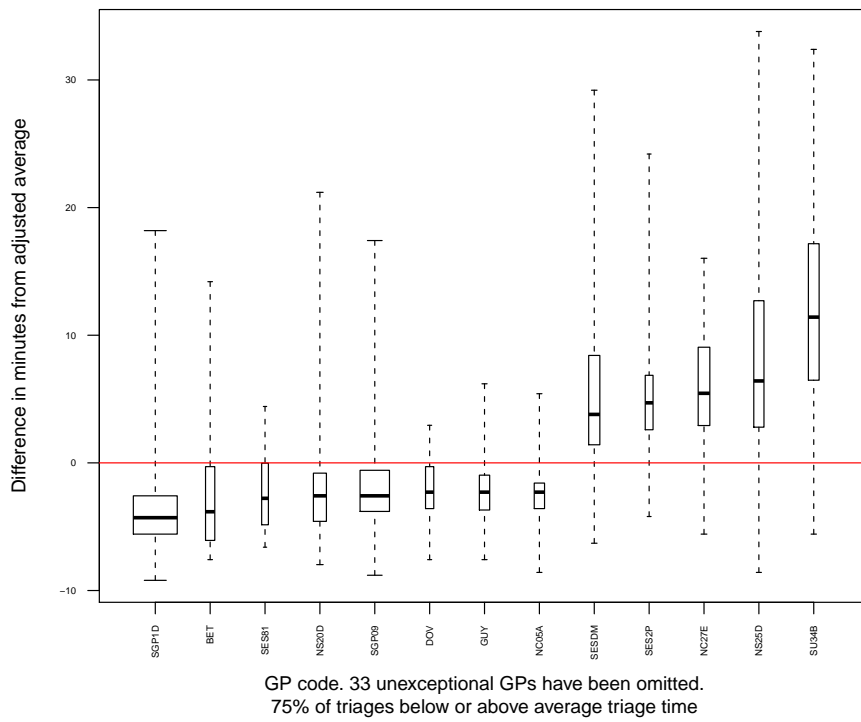


Figure 6.6: Comparing GP triage lengths: Weekday evenings

Saturday morning triages: differences in GP behaviour; 20+ cases per GP

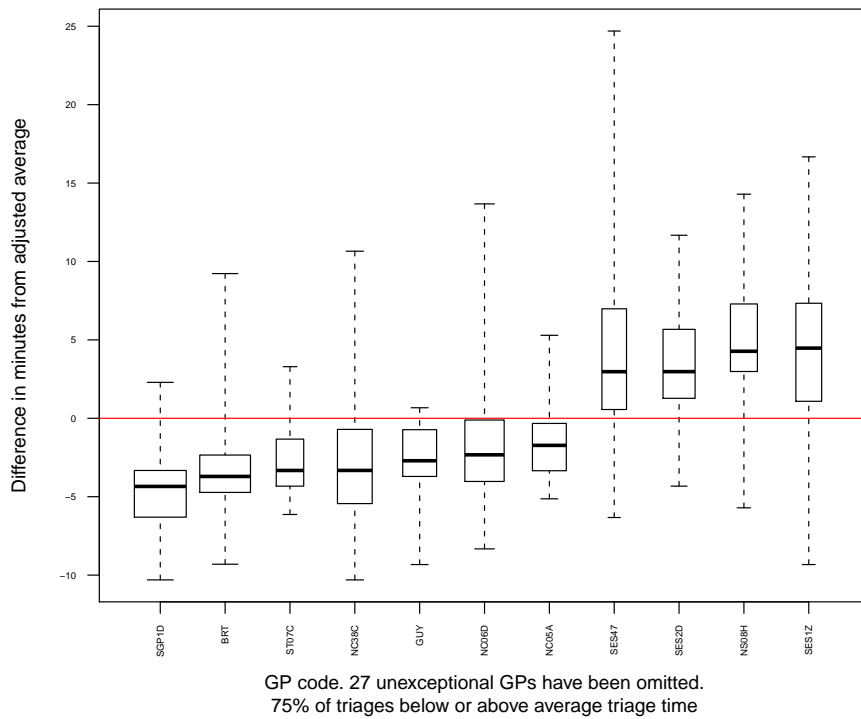


Figure 6.7: Comparing GP triage lengths: Saturday mornings

Chapter 7

Dissemination of Techniques using RExcel

Chapter 7 explains the motivation and implementation of the R package RExcel, used to ensure NDUC staff could use generate forecasts after project closure.

7.1 Motivation

Although R is well built for performing complex statistical analyses, it is not easily accessible to a novice user due to the command prompt, and necessary knowledge of the S language. In reality, not all organisations will have the scope, or need, to employ a statistician full time, so it would be desirable to implement statistical procedures which have the flexibility of R, while still being accessible to the casual user. Baier and Neuwirth [8] discuss this need a little further, and suggest a process of integrating R and Microsoft Excel.

This work was particularly influential on this project, as NDUC continued to use the forecasting procedures in the absence of a statistician. The solution was to use the Excel add-in RExcel, via the R package RExcelInstaller [19]. This allows the user to transfer data between R and Excel, and use R functions in formatted worksheets in Excel, all by the click of a mouse, rather than through a programming language.

7.1.1 Installation and startup

RExcel requires the package `rcom` to be installed first, followed by `statconnDCOM`. Using a working internet connection you can install RExcel by typing the following commands into the R Console, and following the prompt boxes for `statconnDCOM` and RExcel.

```
> install.packages("rcom")
> library(rcom)
> installstatconnDCOM()
> install.packages("REExcelInstaller")
> library(REExcelInstaller)
> installREExcel()
```

This command will automatically install the appropriate version of RExcel (2003 or 2007) to your grade of Office. In what follows, the explanation refers to the Add-In for Excel 2007. Once you have installed RExcel, it is very simple to use. On your usual toolbar at the top, you will have a new tab called Add-Ins, as in figure 7.1. This is where RExcel is hidden. To open R and connect your Excel worksheet you simply click on RExcel and choose the command ‘Start R’, as in figure 7.2. A pop up naming `statconnDCOM` will appear, and then disappear, and an

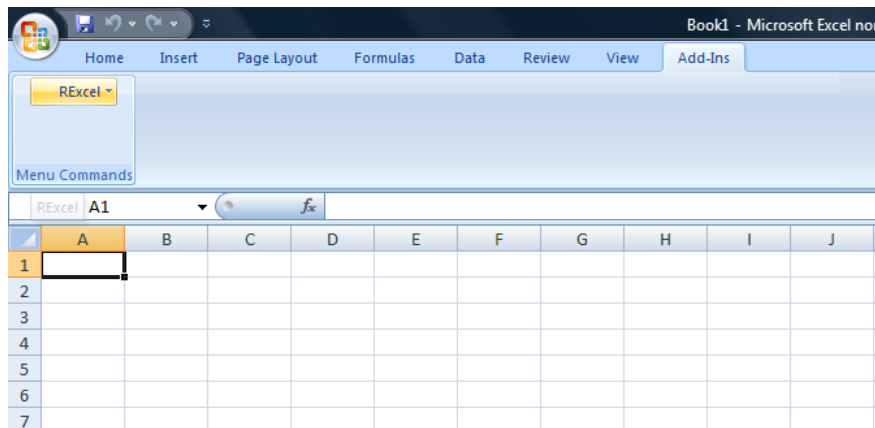


Figure 7.1: Connecting R and Excel: Add-In tab

R Console will open in the background. You are then left with a normal worksheet, connected to R, which can handle R commands as well as Excel functions.

The main commands you can use directly from the Excel spreadsheet, as in figure 7.3, are:

<u>R</u> un code	runs R code from a selected cell(s)
<u>G</u> et R Value	returns R value into that cell
<u>P</u> ut R Var	reads selected cells into R
<u>G</u> et R <u>D</u> ataFrame	returns R dataframe into cells
<u>P</u> ut R <u>D</u> ataFrame	reads selected cells as a dataframe into R

Now we are in a position that we can call the complicated functions we created in R, without the user seeing any of it, while using a program they are familiar with.

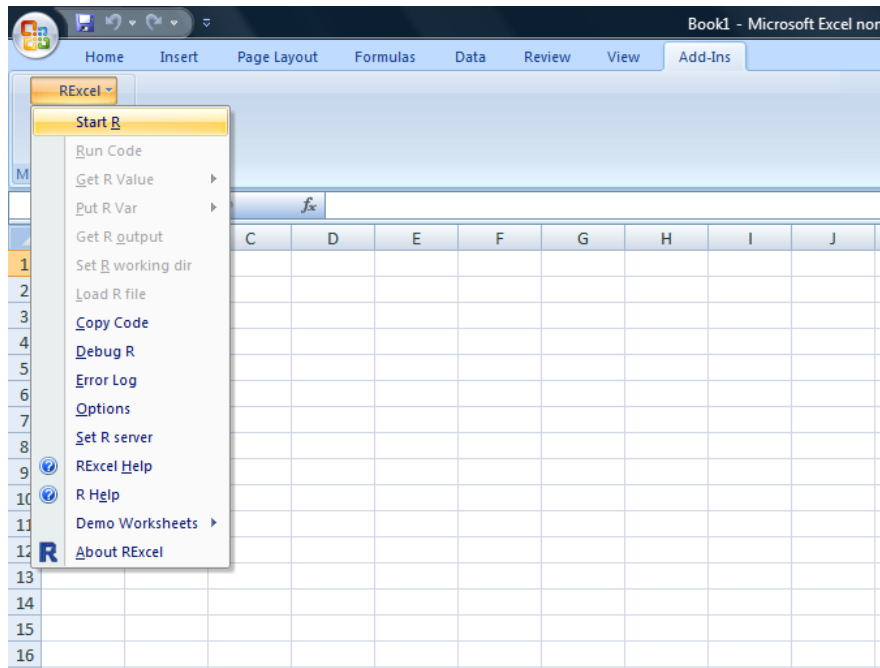


Figure 7.2: Connecting R and Excel: Start R

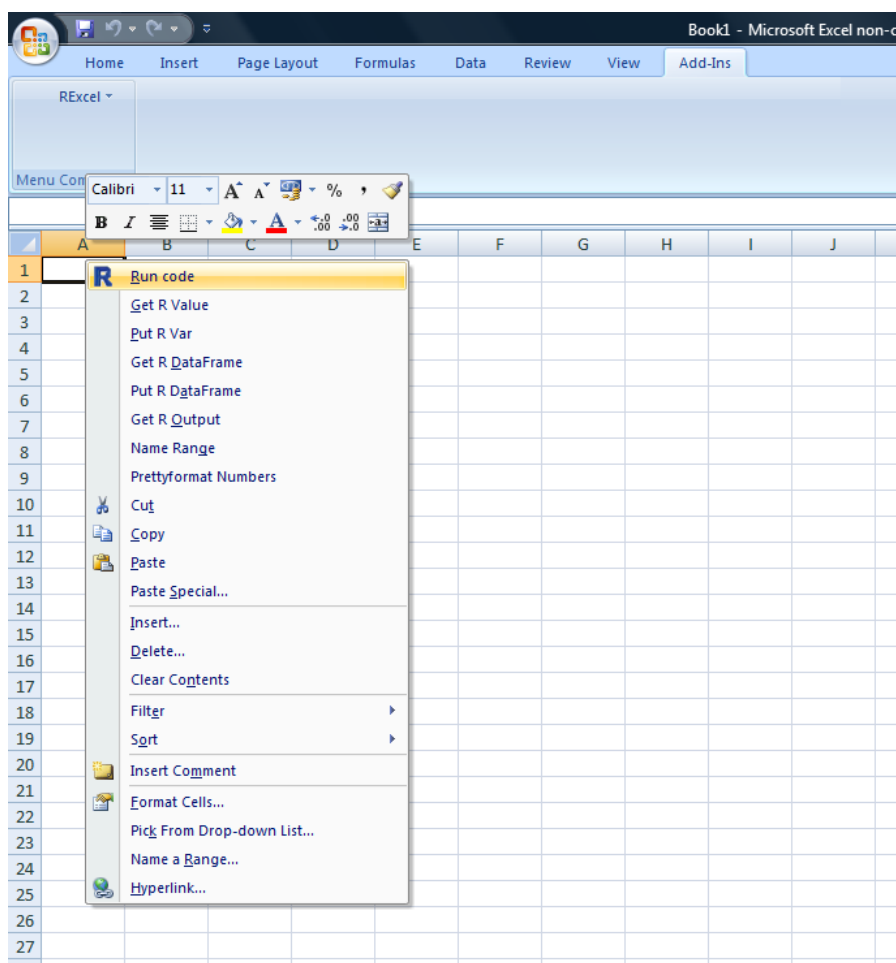


Figure 7.3: Connecting R and Excel: R commands in Excel

August2010 [Compatibility Mode] - Microsoft Excel non-commercial use																
RExcel																
Menu Commands																
C2 01/08/2010 00:02:46																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	urn	taken_by	taken_at	triage_start	triage_finish	cons_start	cons_finish	TakenToTriage	TriageLength	TriageToCons	ConsLength	urgency	UrgencyText	HVUrgency	CVUrgency	area
2	1925790	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	6	6	208	17	9	Routine	Routine		HEX
3	1925791	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	3	11	-11	11	9	Routine			STP
4	1925792	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	0	4	3	14	9	Routine			NGH
5	1925794	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	6	7	-7	7	9	Routine			NGH
6	1925796	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	3	12	68	33	9	Routine	Urgent		HEX
7	1925797	MAU	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	21	7	-7	7	9	Routine			ASH
8	1925798	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	25	9	-9	9	9	Routine			STP
9	1925799	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	10	2	71	6	1	Urgent	Routine		NTP
10	1925800	STO	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	34	6	308	9	9	Routine	Routine		NTP
11	1925801	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	6	5	-5	5	1	Urgent			HEX
12	1925802	MAU	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	38	13	320	19	9	Routine	Routine		NTP
13	1925804	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	59	6	-6	6	9	Routine			STP
14	1925805	MAU	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	80	9	295	14	9	Routine	Routine		NTP
15	1925807	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	81	8	-8	8	9	Routine			STP
16	1925808	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	9	5	41	140	1	Urgent	Urgent		NGH
17	1925809	STO	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	12	7	-7	7	1	Urgent			NGH
18	1925810	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	32	7	346	7	9	Routine	Routine		NGH
19	1925811	STO	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	36	10	-10	10	9	Routine			NGH
20	1925812	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	41	10	-10	10	9	Routine			NTP
21	1925813	MAU	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	48	10	-10	10	9	Routine			NGH
22	1925814	SCMA	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	45	4	-4	4	9	Routine			ASH
23	1925817	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	0	5	35	9	9	Routine			NGH
24	1925818	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	5	8	-8	8	1	Urgent			NTP
25	1925819	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	-8	12	18	2	9	Routine			ALN
26	1925822	MAU	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	6	5	-5	5	9	Routine			STP
27	1925823	STO	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	12	8	-8	8	9	Routine			NGH
28	1925825	IIF	01/08/2010	01/08/2010	01/08/2010	01/08/2010	01/08/2010	0	8	10	11	9	Routine			NTP

Figure 7.4: Raw data extracted using Microsoft Access

7.2 Exception Reporting

Previously, NDUC relied on an historical, monthly, unvalidated performance report. There was no real way of knowing what performance was going to look like for the organisation until the end of the month. Ideally, NDUC wanted a shift by shift performance report, available by 11:00 every morning, so they could highlight any performance issues quickly.

The performance targets were set up to match the monthly report, and source files containing the appropriate KPI functions were created in R to run through RExcel. The first step of the process was to get the raw data into RExcel. You do this by extracting it from an Access database into an Excel spreadsheet. Figure 7.4 shows the screen shot of the raw data. From time to time, there will be blank cells in columns D to K, and it is necessary to replace these blank cells with “#N/A” in order to read the data into R. This is done by using the Excel ‘Find and Replace’ function, which uses the command ‘CTRL H’.

Once you have replaced the blank cells, the worksheet is ready to be read into R. To do this, select all of the data (‘CTRL A’), right click, and choose the option Put R DataFrame (figure 7.6). You will then be asked to name your dataframe, and underneath ‘Dataframe name in R’, type *statstable*, and then click OK (figure 7.7). The R data frame *statstable* now exists in the background of your Excel worksheet.

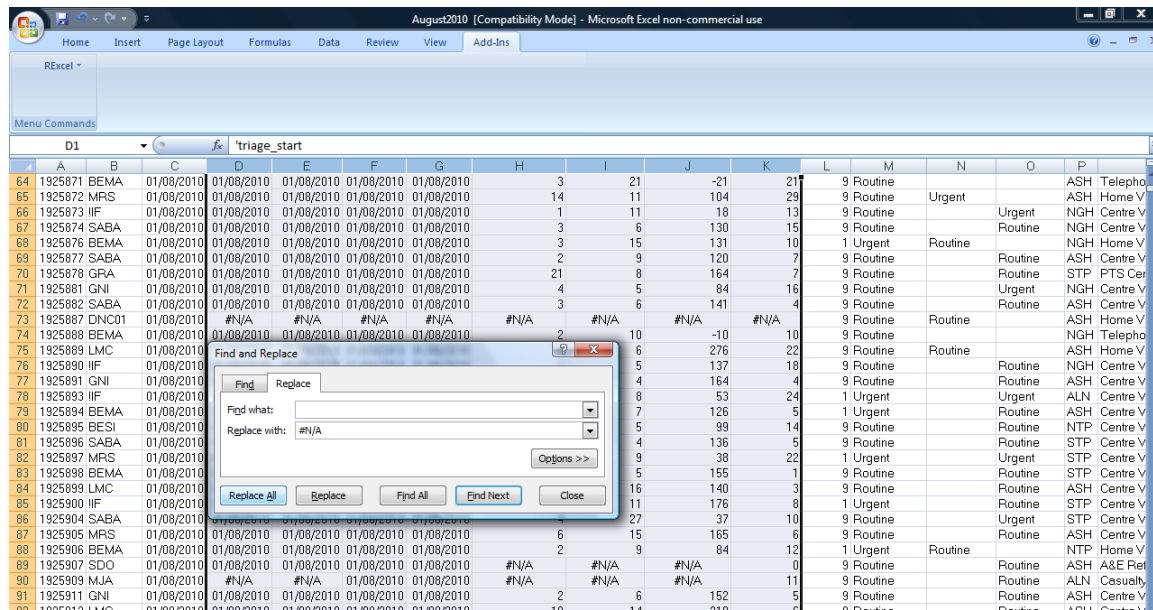


Figure 7.5: Find and Replace in Excel

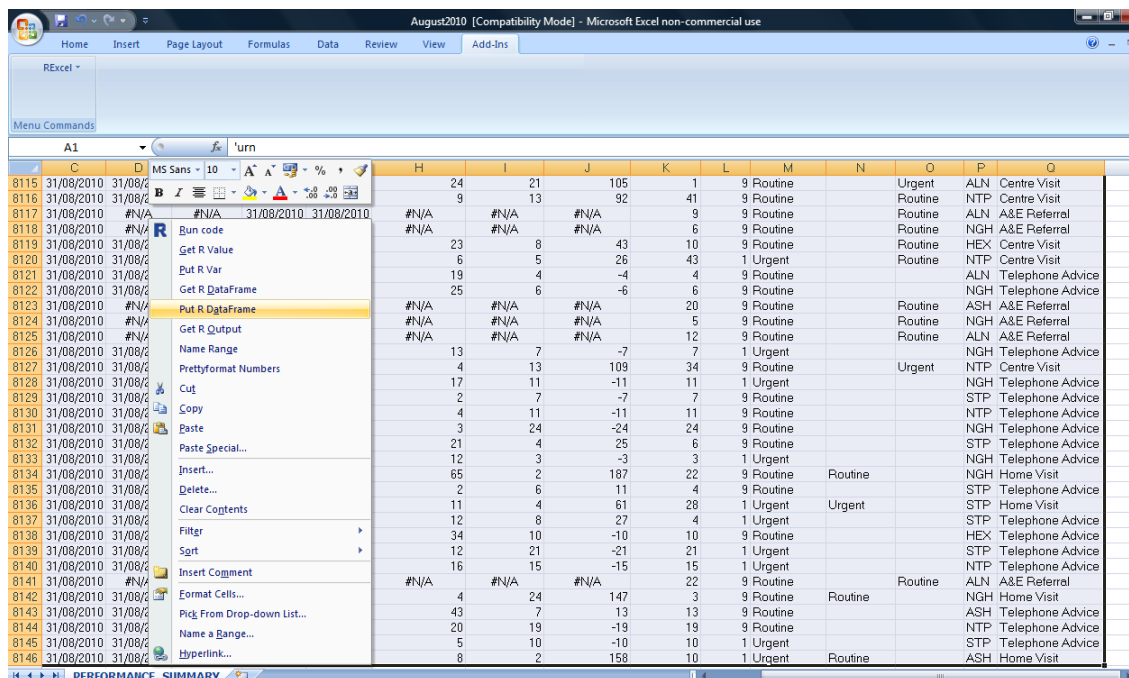


Figure 7.6: Putting the data frame into R: using the command

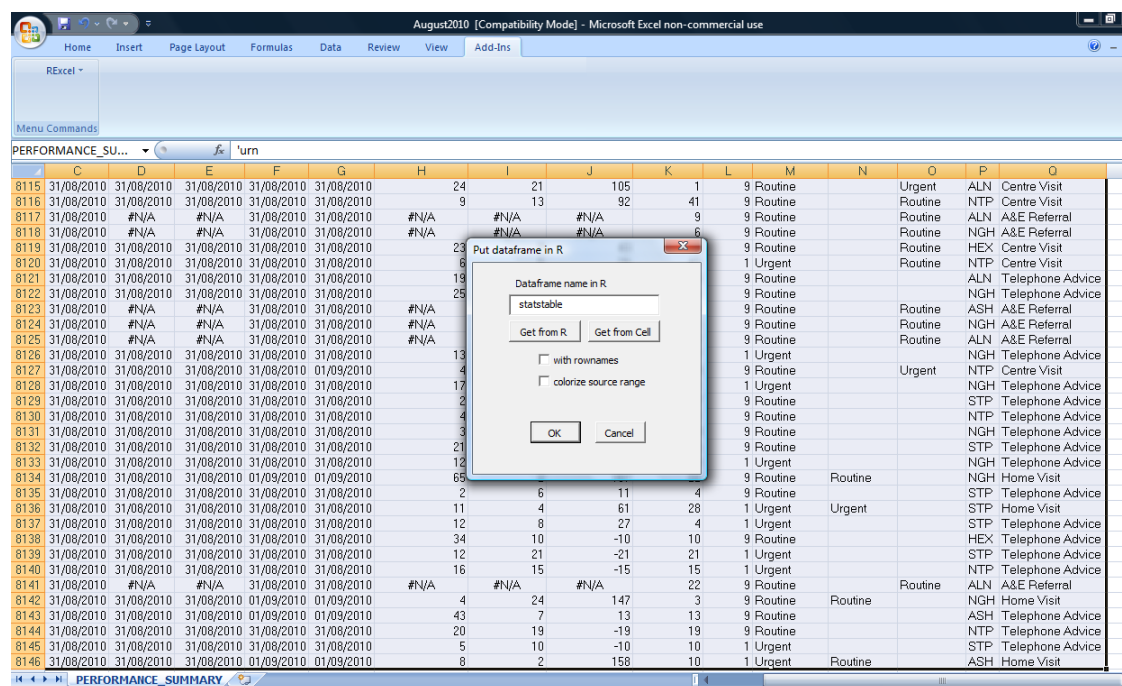


Figure 7.7: Putting the data frame into R: naming the dataset in R

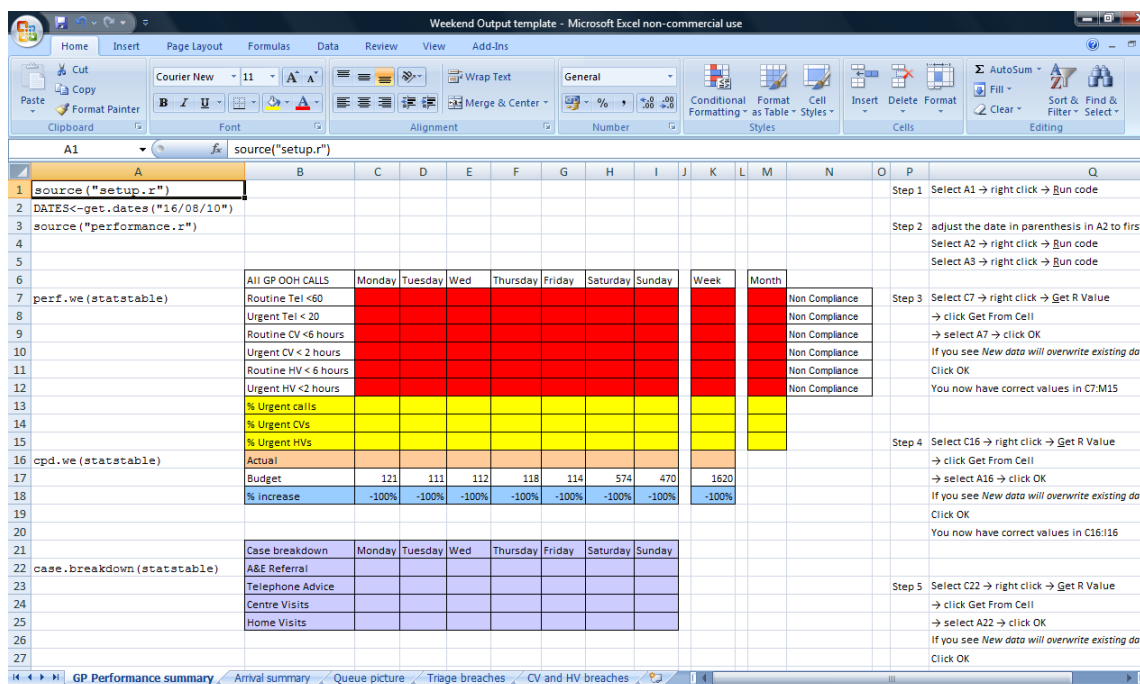


Figure 7.8: Formatted spreadsheet: at a glance

The next stage for the user is to open a pre-formatted worksheet to spit out daily, weekly and month-to-date performance summaries, arrival rate pictures, queue pictures, and a list of breaches for the month. Figure 7.8 shows a screen shot of the worksheet created to produce the weekly performance reports, distributed on Monday mornings. There are five worksheets in total in this document:

1. GP Performance summary: summarises performance over shifts, as well as case breakdown and *CPD*
2. Arrival summary: draws the appropriate *Daily profile* (section 4.1) with the observation overlaid to identify any unexpected peaks or troughs in activity
3. Queue picture: draws the triage queue (section 6.2.3) for the shift
4. Triage breaches: lists all the triage breaches (section 2.4.1) for review
5. CV and HV breaches: lists all the centre visit and home visit breaches (section 2.4.1) for review

Each worksheet has a mixture of R code (always in column A and in **courier new** font), and instructions (generally to the right of the page, and in a different font). Files *setup.r*, *performance.r*, and all *.rdata* files containing the *Daily profiles* have been saved previously to the user's working directory. The steps to complete GP Performance summary are detailed in table 7.1, and are illustrated in figures 7.9 - 7.12. Figure 7.13 shows the completed summary, ready for transfer into a weekly performance report.

Step 1	Select A1 → right click → <u>R</u> un code
Step 2	adjust the date in parenthesis in A2 to first day of the week Select A2 → right click → <u>R</u> un code
Step 3	Select A3 → right click → <u>R</u> un code Select C7 → right click → <u>G</u> et R Value → click Get From Cell → select A7 → click OK If you see <i>New data will overwrite existing data</i> . Click OK to continue Click OK You now have correct values in C7:M15
Step 4	Select C16 → right click → <u>G</u> et R Value → click Get From Cell → select A16 → click OK If you see <i>New data will overwrite existing data</i> . Click OK to continue Click OK You now have correct values in C16:I16
Step 5	Select C22 → right click → <u>G</u> et R Value → click Get From Cell → select A22 → click OK If you see <i>New data will overwrite existing data</i> . Click OK to continue Click OK You now have correct values in C22:I25

Table 7.1: Process for completing GP Performance summary

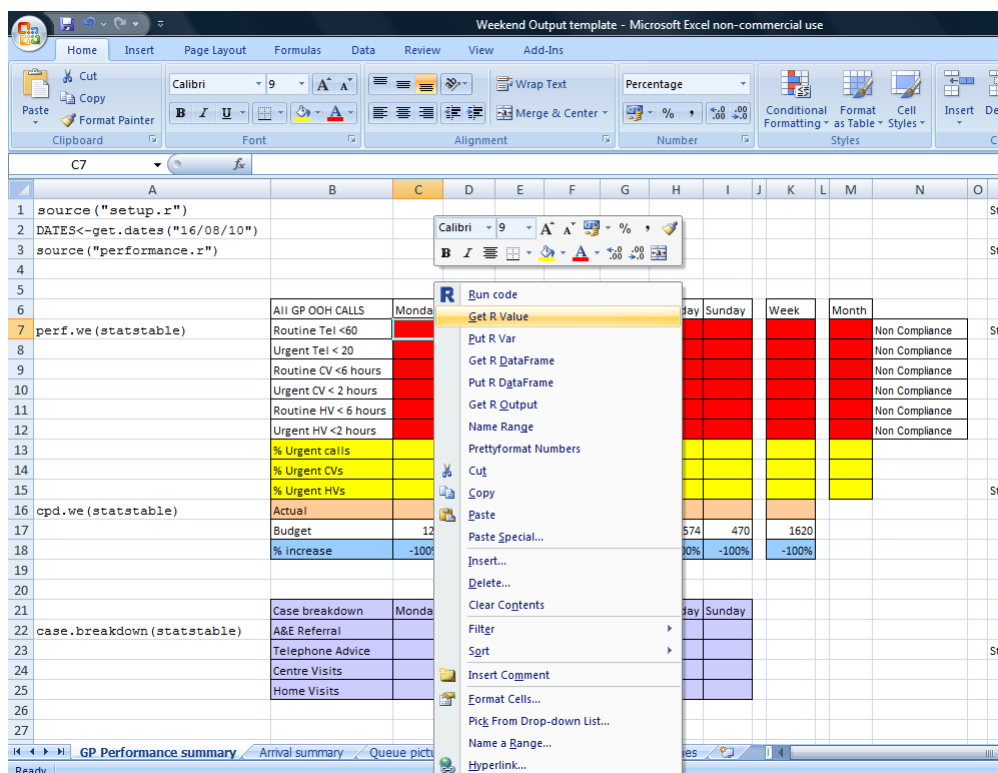


Figure 7.9: Formatted spreadsheet: right click to get R values

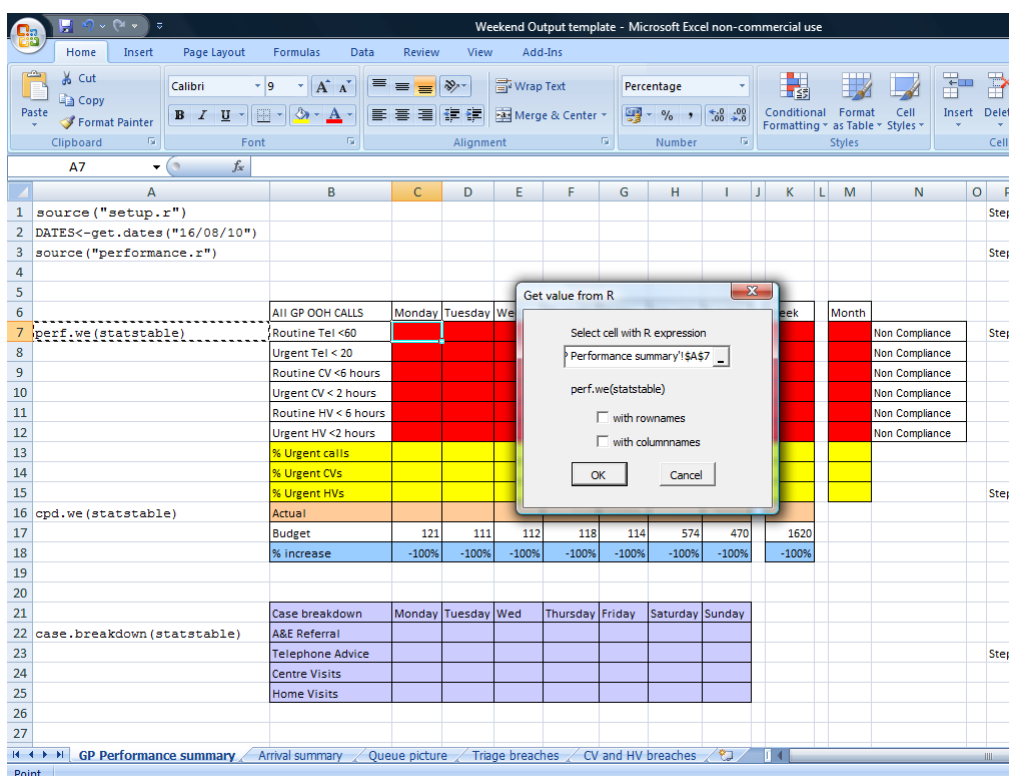


Figure 7.10: Formatted spreadsheet: choosing R code from a cell

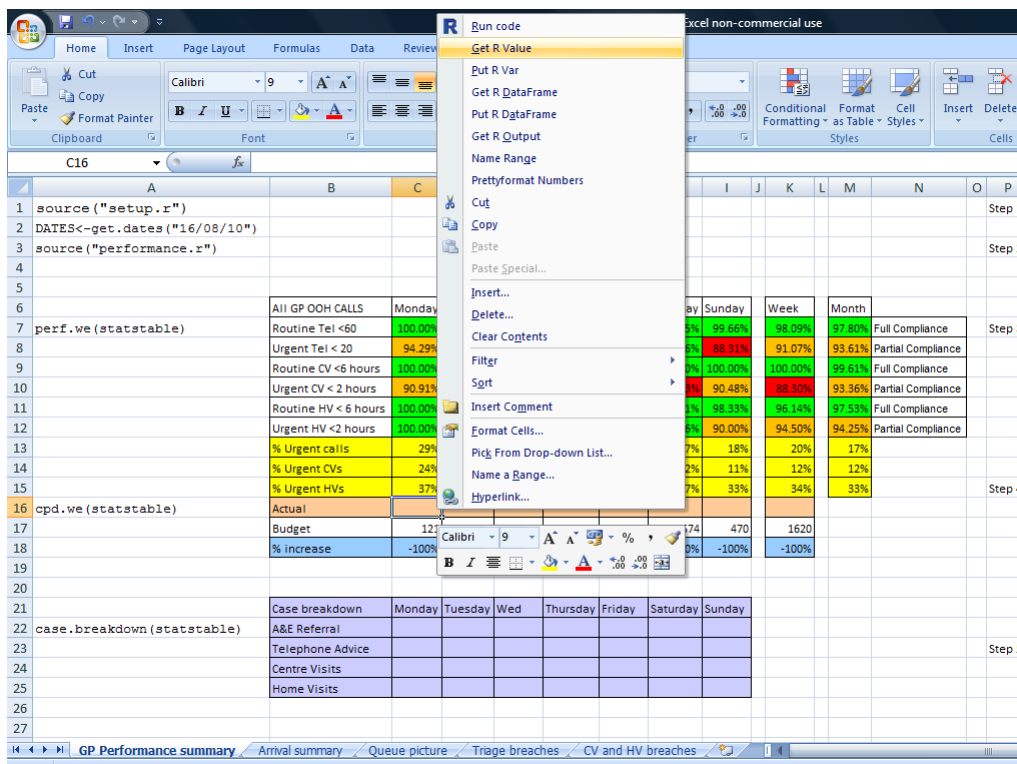


Figure 7.11: Formatted spreadsheet: right click to get R values

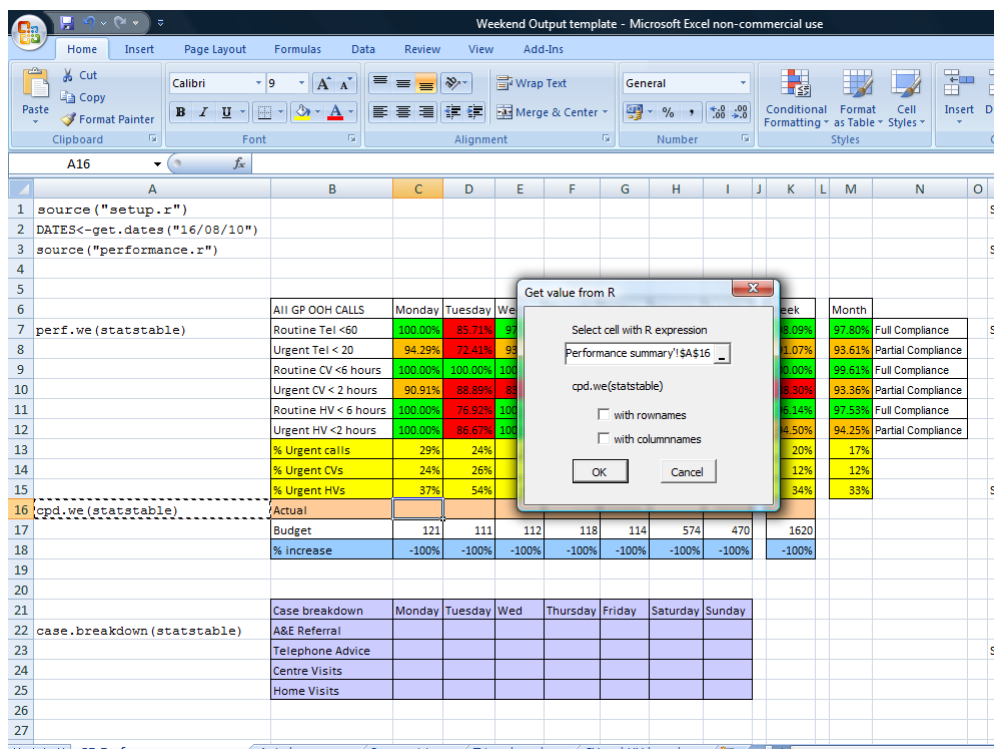


Figure 7.12: Formatted spreadsheet: choosing R code from a cell

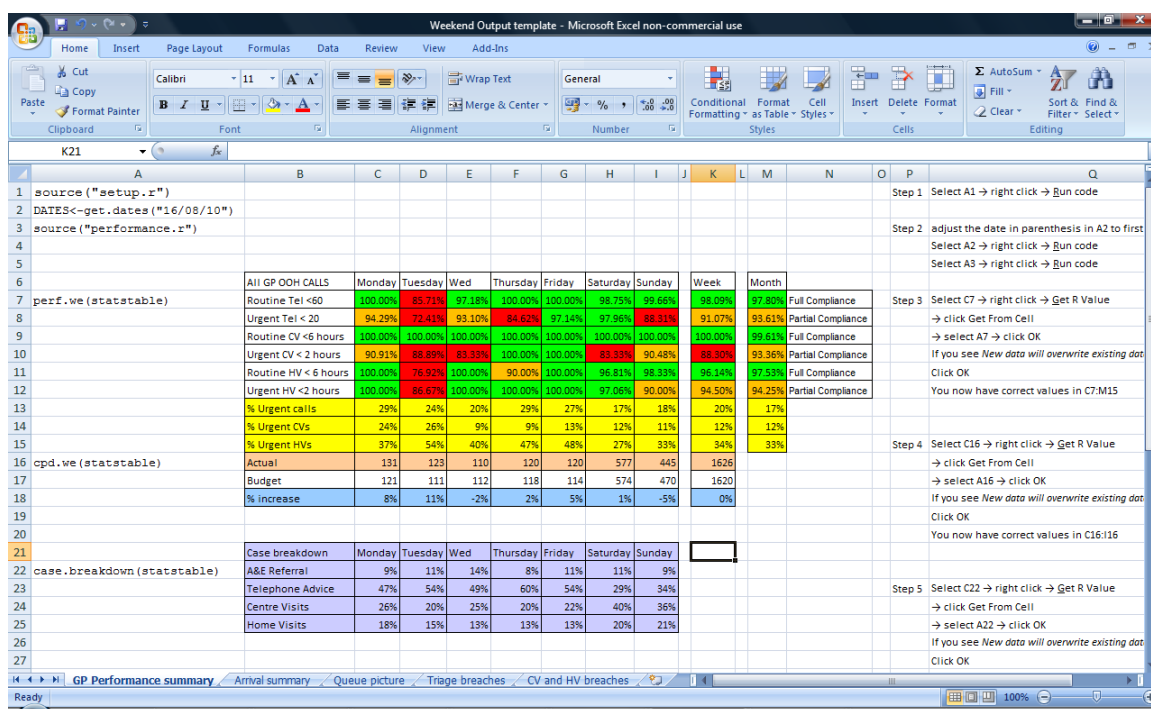


Figure 7.13: GP Performance Summary: completed for week commencing 16/08/2010

Step 1	In cells A1 and A2 make sure the month in the parentheses is correct; january, february, ..., november, december
Step 2	Select A1:A2 → right click → <u>R</u> un code
Step 3	Select A4:A13 → right click → <u>R</u> un code
Step 4	Select A15:A18 → right click → <u>R</u> un code
Step 5	Select A21 → right click → <u>R</u> un code You will now have pdfs of weekday profiles saved in your working directory
Step 6	Select A23 → right click → <u>R</u> un code You will now have pdfs of weekend profiles saved in your working directory

Table 7.2: Arrival summary process

Step 1	Select A7:A20 → right click → <u>R</u> un code
Step 2	Select A23:A29 → right click → <u>R</u> un code You will now have the queue pictures saved in your working directory.

Table 7.3: Queue picture process

Arrival summary and Queue picture worksheets come next. Although these are not intuitively obvious (figures 7.14 and 7.15), they produce and save the R graphics, in .pdf format, with very little effort. The *Daily profile* pictures will be saved as ‘Monday Profile.pdf’,..., ‘Sunday Profile.pdf’, and the queue pictures will be save as ‘Monday.pdf’,..., ‘Sunday.pdf’, all in the user’s working directory. These functions are also designed to read the date into the title, so you can easily see if you’re drawing pictures for the correct week. Arrival summary and Queue picture instructions are detailed in tables 7.2 and 7.3.

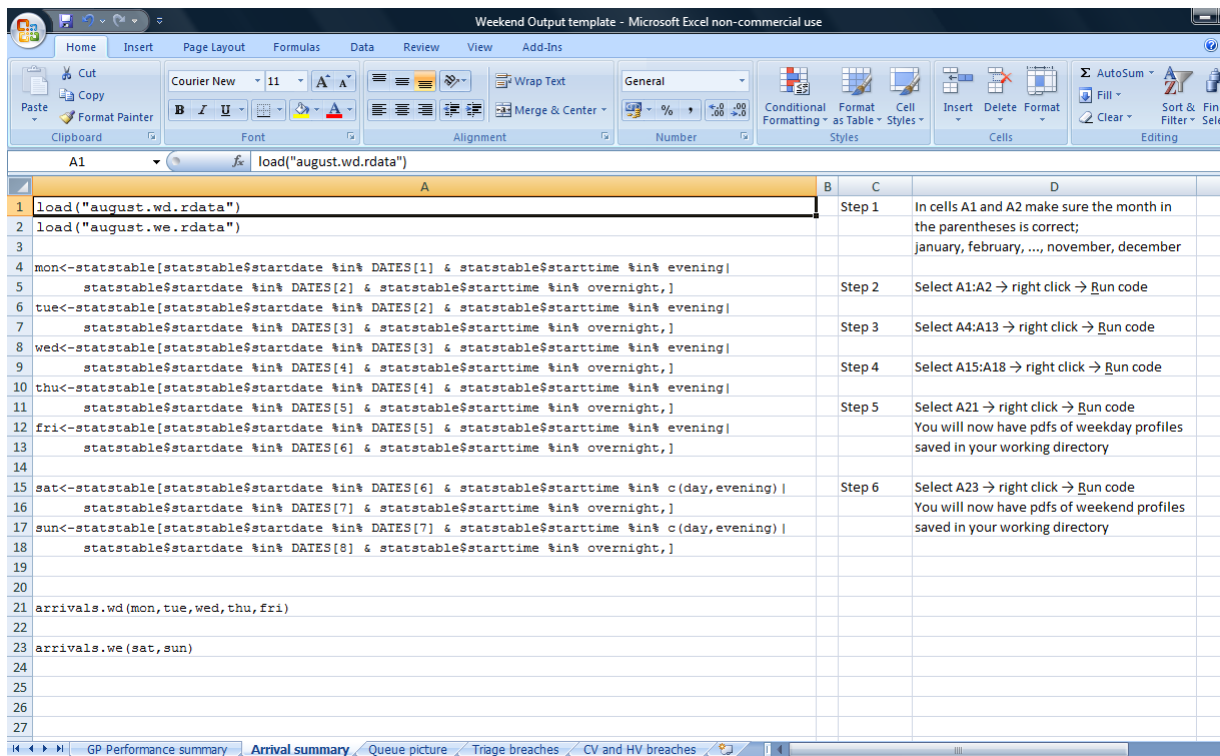


Figure 7.14: Arrival summary: at a glance

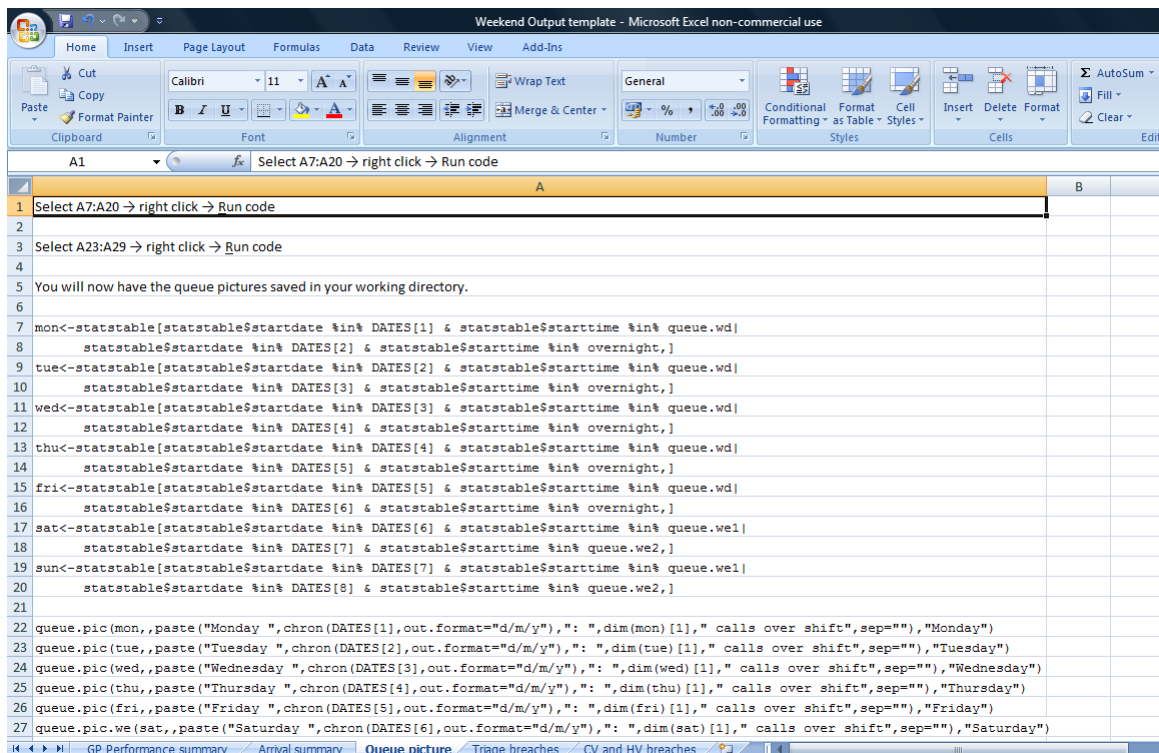


Figure 7.15: Queue pictures: at a glance

Step 1 Select B3 → right click → Get R Value
 →click Get From Cell
 → select A3 → click OK
 You will now have a list of triage breaches
 in front of you
 If 0 appears in cell B3, there were no triage
 breaches

Table 7.4: Extracting breaches

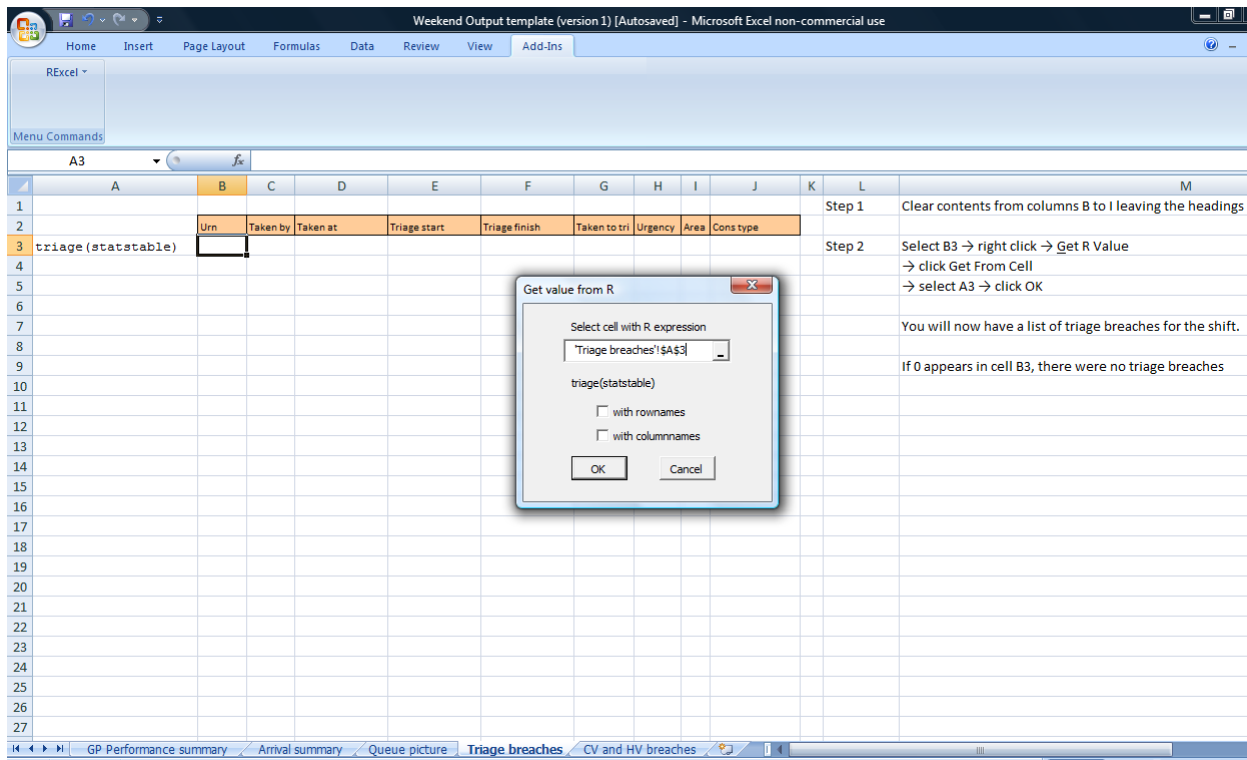


Figure 7.16: Extracting breaches: select R code from cell

The final worksheets generate a full list of breaches to be reviewed by the Operations Manager. The process is, again, very simple and generating the triage breaches is illustrated in figures 7.16 and 7.17. The function you need to use is `triage(statstable)`, and this outputs a full list of triage breaches, and is written in cell A3 of the worksheet. Centre Visit and Home Visit breaches follow the same process, using functions `cv(statstable)` and `hv(statstable)`.

Weekend Output template (version 1) [Autosaved] - Microsoft Excel non-commercial use

Home Insert Page Layout Formulas Data Review View Add-Ins

RExcel

Menu Commands

B3 1925805

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		Urn	Taken by	Taken at	Triage start	Triage finish	Taken to tri	Urgency	Area	Cons type		Step 1	Clear contents from columns B to I leaving the heading
3	trriage (statstable)	1925805	MAU	01/08/2010 01:54	01/08/2010 03:14	01/08/2010 03:23	80	Routine	NTP	Home Visit		Step 2	Select B3 → right click → Get R Value
4		1925807	IIF	01/08/2010 02:02	01/08/2010 03:23	01/08/2010 03:31	81	Routine	STP	Telephone Advice			→ click Get From Cell
5		1926724	JUP	03/08/2010 18:49	03/08/2010 20:27	03/08/2010 20:36	98	Routine	HEX	Telephone Advice			→ select A3 → click OK
6		1926825	IIF	04/08/2010 02:30	04/08/2010 03:47	04/08/2010 03:57	77	Routine	STP	Home Visit			
7		1926826	MAU	04/08/2010 02:42	04/08/2010 04:29	04/08/2010 04:34	107	Routine	NTP	Home Visit			You will now have a list of triage breaches for the shift
8		1926830	MAU	04/08/2010 03:27	04/08/2010 04:52	04/08/2010 04:56	85	Routine	NGH	Telephone Advice			
9		1926838	MAU	04/08/2010 06:10	04/08/2010 07:33	04/08/2010 07:46	83	Routine	NTP	Telephone Advice			If 0 appears in cell B3, there were no triage breaches
10		1927223	BESI	05/08/2010 19:26	05/08/2010 20:37	05/08/2010 20:40	71	Routine	STP	Home Visit			
11		1927455	JOY	06/08/2010 19:05	06/08/2010 20:21	06/08/2010 20:36	76	Routine	ALN	Telephone Advice			
12		1927554	BEMA	06/08/2010 23:17	07/08/2010 00:21	07/08/2010 00:30	64	Routine	ASH	Telephone Advice			
13		1927569	SCP	07/08/2010 01:23	07/08/2010 02:29	07/08/2010 02:40	66	Routine	STP	Home Visit			
14		1927570	MAU	07/08/2010 01:27	07/08/2010 02:40	07/08/2010 02:50	73	Routine	ASH	Telephone Advice			
15		1927572	LMI	07/08/2010 01:35	07/08/2010 02:50	07/08/2010 02:56	75	Routine	ASH	Telephone Advice			
16		1927574	IIF	07/08/2010 01:46	07/08/2010 02:58	07/08/2010 03:08	73	Routine	ASH	Telephone Advice			
17		1927576	LMI	07/08/2010 02:09	07/08/2010 03:21	07/08/2010 03:33	72	Routine	ASH	Telephone Advice			
18		1927585	LMI	07/08/2010 03:57	07/08/2010 05:12	07/08/2010 05:34	75	Routine	NGH	Telephone Advice			
19		1927588	SCP	07/08/2010 04:15	07/08/2010 05:22	07/08/2010 05:42	67	Routine	NTP	Home Visit			
20		1927616	AHA	07/08/2010 07:48	07/08/2010 09:02	07/08/2010 09:08	74	Routine	NTP	Telephone Advice			
21		1927990	OLD	07/08/2010 12:43	07/08/2010 14:02	07/08/2010 14:40	79	Routine	NGH	Telephone Advice			
22		1928283	TRHA	07/08/2010 19:12	07/08/2010 20:46	07/08/2010 20:58	94	Routine	STP	Centre Visit			
23		1928589	BEMA	08/08/2010 10:44	08/08/2010 11:48	08/08/2010 12:02	63	Routine	NGH	Telephone Advice			
24		1928602	GCP	08/08/2010 10:54	08/08/2010 11:57	08/08/2010 12:07	63	Routine	NGH	Telephone Advice			
25		1928850	KADI	08/08/2010 17:12	08/08/2010 18:39	08/08/2010 18:46	87	Routine	HEX	Telephone Advice			
26		1930063	JOY	13/08/2010 18:53	13/08/2010 20:03	13/08/2010 20:11	70	Routine	NTP	Centre Visit			
27		1930112	ROD	13/08/2010 20:58	13/08/2010 23:04	13/08/2010 23:05	126	Routine	ALN	Home Visit			

GP Performance summary Arrival summary Queue picture Triage breaches CV and HV breaches

Figure 7.17: Extracting breaches: completed for August to date

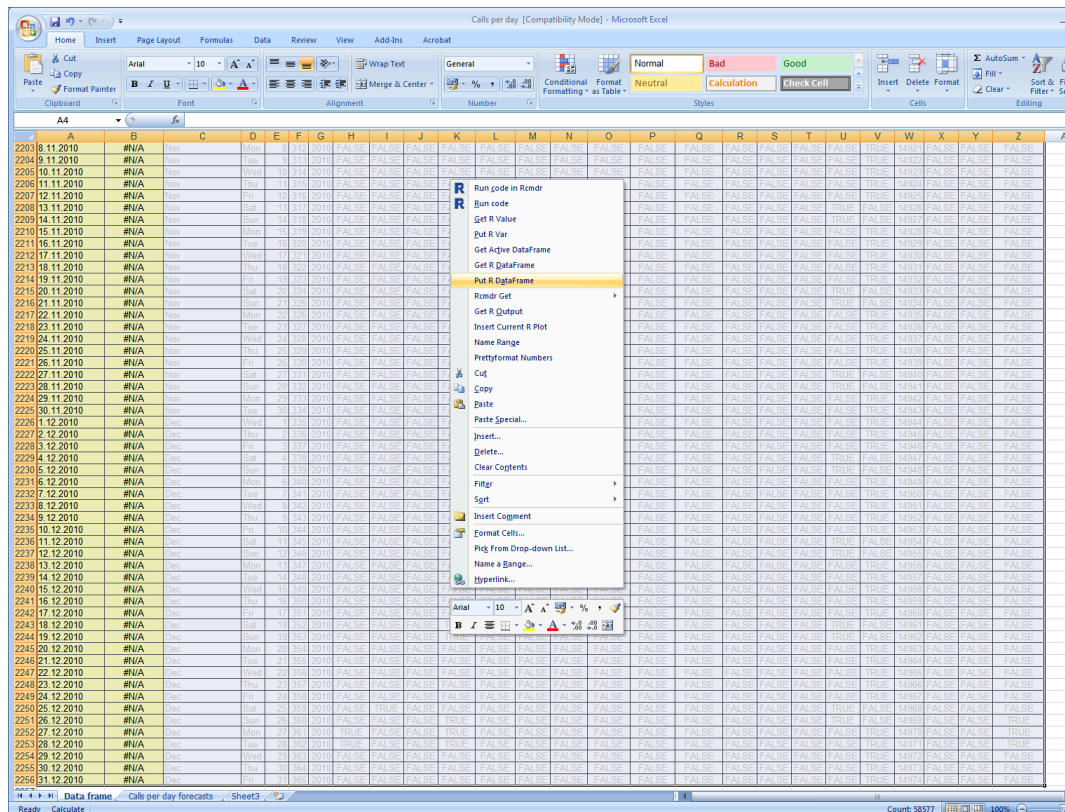
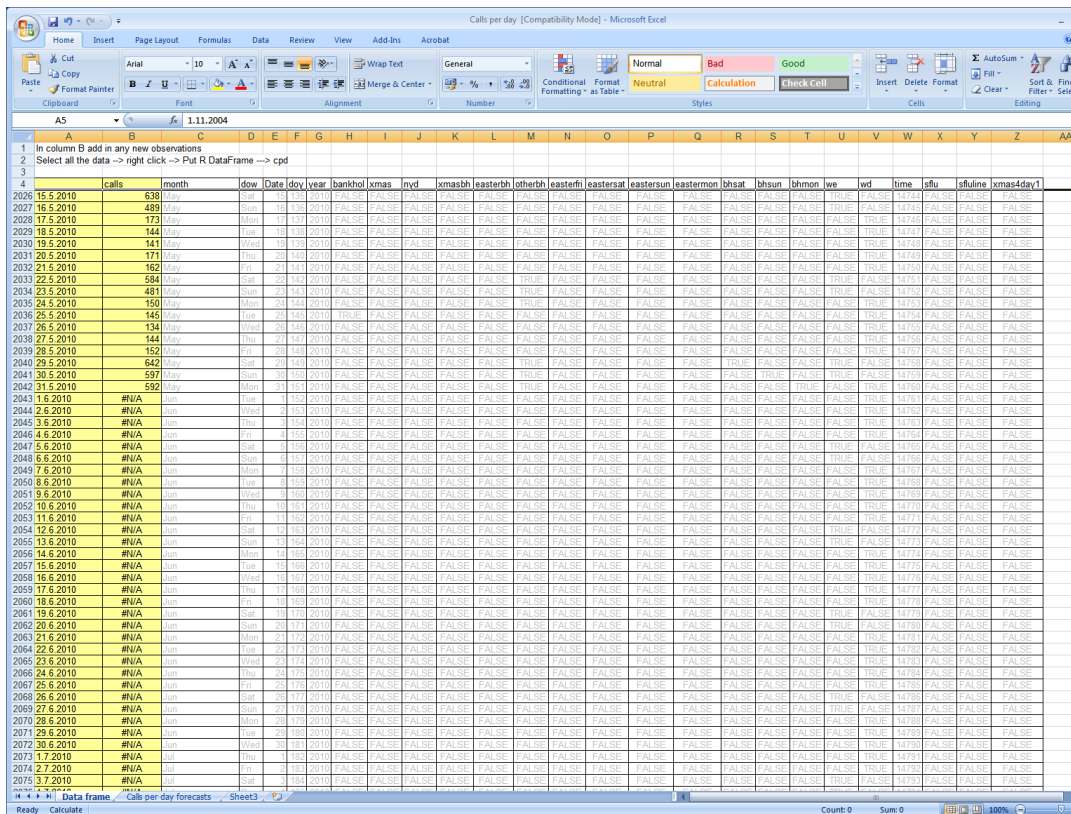
7.3 Transferring *CPD* model into RExcel

Recall chapter 3 which described how we forecast calls per day *CPD*, and recall the linear model in equation 3.21:

$$\begin{aligned}
 CPD = & \beta_0 + \beta_1dow + \beta_2month + \beta_3year + \beta_4time + \\
 & + \beta_5easterfri + \beta_6eastersat + \beta_7eastersun + \beta_8eastermon + \\
 & + \beta_9bhsat + \beta_{10}bh sun + \beta_{11}bhmon + \\
 & + \beta_{12}maybh1 + \\
 & + \beta_{13}xmasbhwd + \beta_{14}xmasbhwe + \\
 & + \beta_{15}nydwd + \beta_{16}xmaswd + \\
 & + \beta_{17}xmaswe + \beta_{18}nydwe + \\
 & + \beta_{19}xmas4day1 + \\
 & + \beta_{20}month : we + \beta_{21}dow : xmas4day1 + \\
 & + \beta_{22}month : year + \beta_{23}month : time + \\
 & + \beta_{24}time : easterbh + \\
 & + \beta_{25}sflu + \beta_{26}sfluline + \\
 & + \beta_{27}sflu : sfluline + \epsilon
 \end{aligned} \tag{7.1}$$

To transfer this into RExcel we created two separate Excel worksheets to pilot among staff at NDUC. The first step is to setup the data frame of previous observations, and the second is to declare the linear model *CPD* in R, and generate forecasts. Setting up the data can be done using an Excel spreadsheet (worksheet ‘Data frame’) as illustrated in figure 7.18.

Column B exists for the user to input new observations, and is coloured in pale yellow. The rest of the worksheet does not need to be touched, and contains the fixed covariates used in equation 7.1. As before, you need to read this data into R before you continue. Once new observations have been included, select the data, right click, choose ‘Put R DataFrame’, and name it *cpd*. The data frame *cpd* now exists in the background of your Excel worksheet. This is illustrated in figure 7.19.



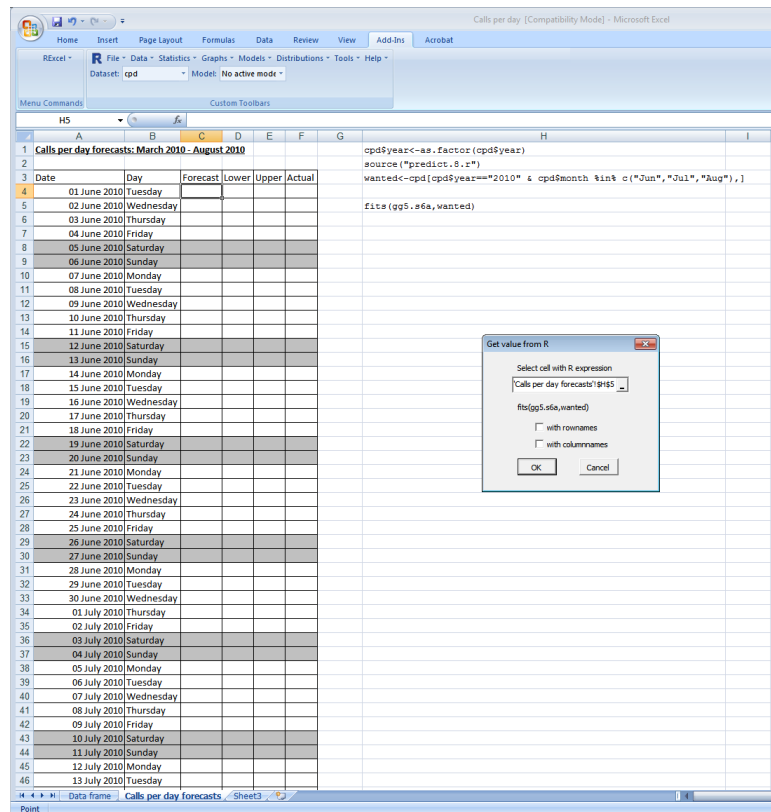


Figure 7.20: Calls per day: Calls per day Forecast worksheet

Now move onto the 'Calls per day forecasts' worksheet. In this illustration we have data up until 31/05/2010, and are forecasting *CPD* for June, July and August. The file *predict.r* has been saved to the user's working directory. Choose cells H1:H3, right click, and choose 'Run code'. Then in cell C4, right click, and choose 'Get R Value', and then 'Get From Cell'. Choose cell H5 which contains the function *fits*, and new *CPD* forecasts with upper and lower limits will appear in columns C, D, and E. This process is illustrated in figures 7.20 and 7.21.

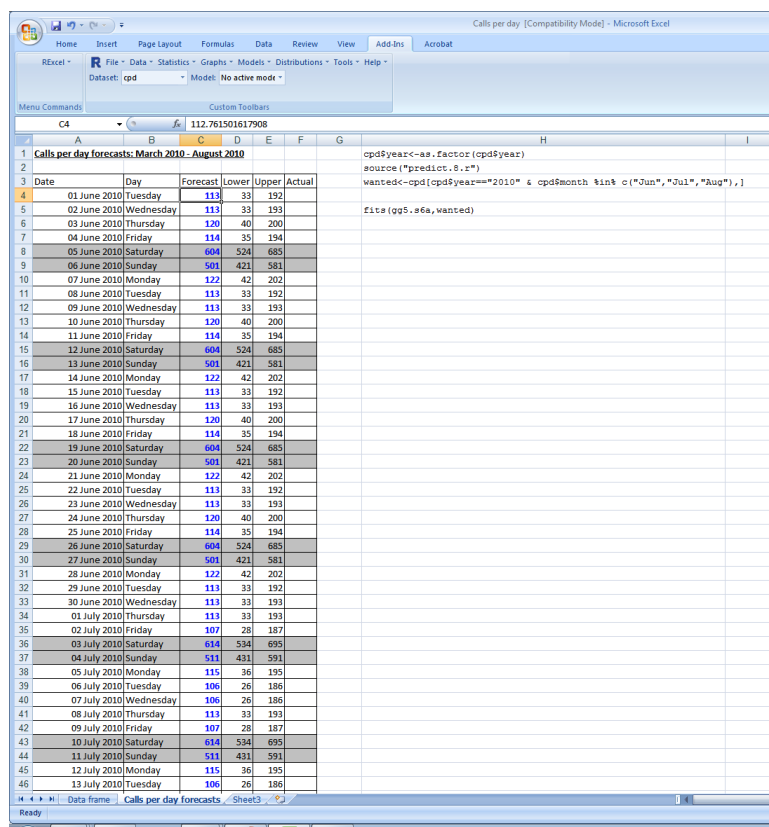


Figure 7.21: Calls per day: Updated forecasts

7.4 Conclusions and Recommendations

This approach was adopted and used by members of the IT department from July 2010, with a handover period of about 3 months. User feedback was good, with the process accessible to IT and Analyst level.

In practice, these worksheets will only be as good as the models and the users. Users will not know how to adjust the code behind these worksheets, and a statistician will need to be consulted to fix any issues. There is also the ‘garbage in - garbage out’ risk. The method is quite specific, particularly when naming your datasets. If you named the data range *stastable*, rather than *statstable*, the process would not work.

The GP activity plots (section 6.2.2) can also be drawn using RExcel, by extracting the data from the Adastra database through MySQL Server, and following steps similar to before. This stage has been omitted from this chapter as there is still a manual element of the process; cross-referencing GP names with a program RotaMaster to see what shifts they were working. Further work would connect the two sources of data, Adastra and RotaMaster, but in the meantime, the principals of generating the GP activity plots in RExcel are identical to the Queue picture process.

In the absence of someone well versed in R, I would not recommend transferring the boxplot monitoring method (section 6.2.4) into RExcel, as it is too complicated. Once this process has been set up, however, it should not take an expert more than one day to run, therefore companies who cannot afford a full time statistician could employ one only when they need them.

Chapter 8

Further work

Before we conclude this thesis, it is appropriate to mention further work that could be carried out to address a number of other issues relating to this work, in particular, more detailed modelling. This chapter discusses two important issues around out-of-hours call centres; queuing and government legislation, before making some suggestions on how this work could be carried forward.

8.1 Queues

Recall figure 2.2 in section 2.4.1. For performance measurements we need to know how many urgent and routine patients are at server 1, server 2, and server 3.

Figure 8.1 shows nine potential patient paths. Let t_0 be the time a patient enters the triage queue. This patient has already spoken to a call handler, been assessed as urgent or routine, and been told a GP will ring them back when they are available. Let U_W and R_W be the time, in minutes, the patient waits for the GP to call them back for an urgent and routine case respectively. To meet the KPI we need 95% of U_W and R_W to obey the following:

$$U_W \in [0, 20]$$

$$R_W \in [0, 60]$$

The next step is triage by a GP. This is GP dependent. Let T_U and T_R be the length of time, in minutes, the GP spends triaging an urgent or routine call, respectively. There are no national standards to report on this, but as an organisation we performance-manage this by expecting our GPs to triage at least five cases per hour. That means:

$$T_U < 12$$

$$T_R < 12$$

If the patient's ailment has been dealt with over the phone, they leave the system, and their case is recorded as Telephone Advice. If there is further assessment and the GP arranges a face-to-face consultation, the patient enters a second queue. Let HU_W and HR_W be the time, in minutes, a patient has to wait between the end of the triage and the start of a home visit consultation for urgent and routine cases respectively. Similarly, let CU_W and CR_W be the time, in minutes, a patient has to wait between the arrival at a UCC and the start of a centre visit consultation for urgent and routine cases respectively. To meet the KPI we need 95% of HU_W , HR_W , CU_W and CR_W to obey the following:

$$HU_W \in [0, 120]$$

$$HR_W \in [0, 360]$$

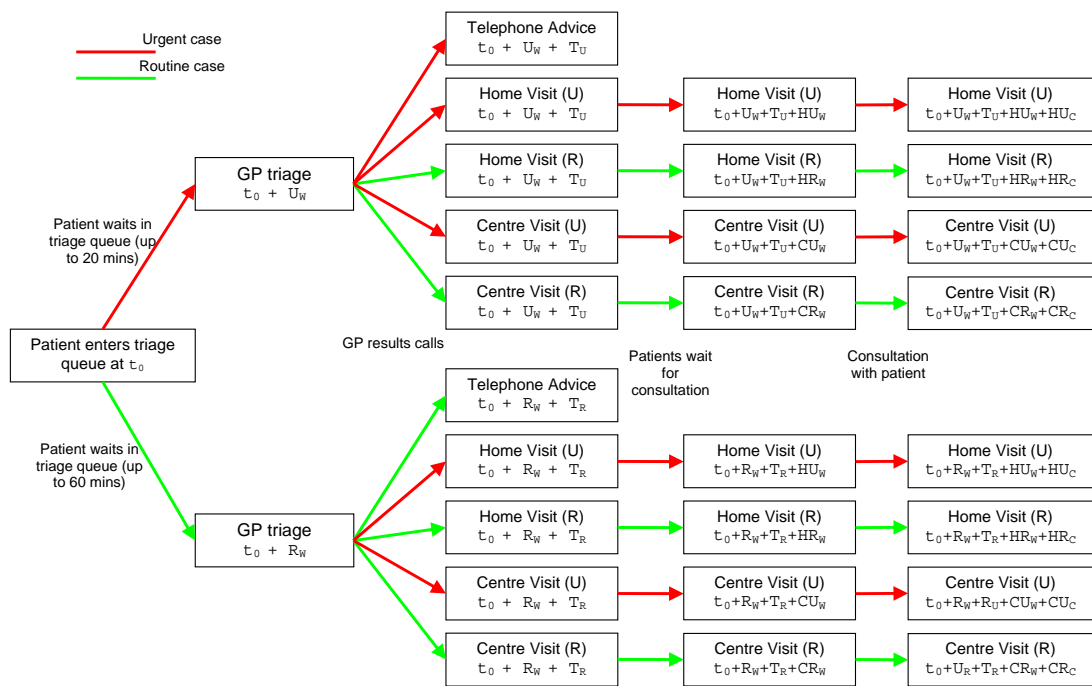


Figure 8.1: Nine potential patient paths

$$CU_W \in [0, 120]$$

$$CR_W \in [0, 360]$$

Let HU_C and HR_C be the time, in minutes, of the GP home visit consultation urgent and routine cases respectively. Similarly, let CU_C and CR_C be the time, in minutes, of the GP consultation at the UCC, for urgent and routine cases respectively.

Using the working described in this thesis around forecasting arrival rates, and GP triage length distributions, a simple extension would be to use this, and the above description to simulate queues at NDUC.

We could link GP activity to queue lengths, as mentioned in section 6.2.3. From this we can build an understanding of how queues develop, and how long it takes to service them. Then we can assess how best to plan GP shifts, and other closely related workers, like drivers (recall a gap in resource between drivers and GPs in section 2.5.2). Furthermore, we can develop work we did not touch on in this thesis, but that is important to NDUC. How can we ensure any patient using the service is receiving the correct method of healthcare? Currently we send out patient questionnaires, as we are required to by our commissioners. Could there be a better way to evaluate this patient feedback? Could we design a better method of capturing this feedback? All of these will answer questions about patient pathways through GP out-of-hours, which is likely to be extremely useful in the wider healthcare system.

8.2 Government vision for the future of the NHS

The recent White Paper [6] caused a stir in among the health community, and in particular, GP out-of-hours providers and commissioners. Urgent Health UK Ltd is a federation of social enterprise unscheduled care providers [1], and NDUC is a member of this organisation. Urgent Health UK (UHUK) is currently made up of nine leading out-of-hours and unscheduled care providers who together serve over 9 million people and have a 17% share of the out-of-hours service in England and Wales. The purpose of UHUK is to facilitate a forum for several out-of-hours providers so that best practice can be adopted. It is quite an interesting idea as there may be times when the member organisations are competing for contracts. The annual UHUK conference took place in Birmingham this year and the White Paper was discussed in some detail by the providers, raising concerns mentioned below.

8.2.1 GP consortia

Currently, Primary Care Trusts (PCTs) exist, and are responsible for local health improvement, and commission out-of-hours providers, such as NDUC. PCTs are generally made up of non-clinicians, and manage several different patient services in their local area. The process of awarding contracts often goes through lengthy Pre Qualification Questionnaire (PQQs) or Invitation To Tender (ITTs), before interviews and announcing preferred bidders. The vision is to scrap PCTs completely, and replace them, instead, with GP consortia, who will commission out-of-hours services.

This in itself has raised concerns among both providers, like NDUC, and commissioners, the PCTs. If GP consortia are commissioning, this creates a situation where commissioners can also be providers. What is to stop GP consortia deciding a provider is too expensive, and then setting themselves up as a new provider, and undercutting the competition? On the other hand, who is better to lead the out-of-hours commissioning process? GPs work in local surgeries seeing their patients all the time, and know their needs.

If the vision of the White Paper goes ahead, it could completely change the current GP out-of-hours system, and our project would need to be reassessed for relevance.

8.2.2 Single Point of Access

The second issue for out-of-hours providers is the introduction of the Single Point of Access (SPA), through the provision of a new three digit number, 111, to call when “it’s less urgent than 999”. This is currently being piloted in Durham and Darlington and Dr Katherine Noble, a GP and Clinical Lead for 111 in Durham and Darlington, spoke at the UHUK conference giving an update on the service.

Before the pilot in the North East of England, there was a lot of research done around current GP out-of-hours, and the choice of the SPA number. A WHICH? report at the end of 2008 suggested that 52% of the population did not even know there was a GP out-of-hours service. This, I could naturally believe, as before I started this KTP project I did not know there was a GP out-of-hours service, and most people I have spoken to over the last two years did not know this either. There were focus groups about trying to come up with the right SPA number as well. Even those that are aware of the GP out-of-hours service said that in an urgent, but non-emergency situation, 999 is the only number they can remember. NDUC’s number is 0845 60 80 320, and if you need to use the service you would probably have to check the internet to remind you of the number. Therefore, the choice of 111 seemed natural.

Future work will have to take into consideration the implementation of SPA. It will directly affect call volumes of GP out-of-hours, and will need to be factored into forecasting models when it goes nationwide.

Chapter 9

Conclusions

During this research, we spent time familiarising ourselves with the, somewhat unknown, GP out-of-hours service, and the requirements providers face from their commissioners. We built a framework of a simple patient path, and highlighted that GP resource is, by far, the most expensive when running such a service. Therefore, all further work was centred on the activity the GPs would have to deal with, and indeed, the activity of the GPs themselves.

Two different forecasting methods were implemented at NDUC to predict calls per day, and arrivals rates into a call centre providing GP out-of-hours services to the North East of England. This was done through quite simple methods of linear regression and locally weighted scatterplot smoothing. In practice, these simple methods worked, and were well received by the company, so we continued to develop these for the final implementation. They are, also, by no means perfect, but work well in this application, and further investigation could improve them even more. Particular attention needs to be paid to time periods around Christmas.

This was certainly not the end of our story, and several interesting problems arose throughout development. What do we do when a pandemic arises and throws the *CPD* and *Daily profile* models out the window? We faced the issues of this in the middle of our project, and felt the pressures from our commissioners and our patients, tackling it with updated forecasting methods through moving averages. Time was also a pressure here and we weren't able to implement the updated forecasting procedures until October 2009. Documentation of the process in October now exists, and should be transferrable to another pandemic situation, ensuring the healthcare system does not get hit the same way it did in July 2009.

In contrast to forecasting, we investigated the sensitive problem of monitoring GP triage lengths when working out-of-hours, using particular exclusions, comparing 'like with like' as far as possible, and adopting a generalised linear model. This took the longest time during our research as we wanted to produce something that could be used by any out-of-hours organisations using Adastra. We also knew that GPs would question every step of the monitoring process, so getting that right was key. This work was presented to a group of out-of-hours GPs, with an explanation of the methodology, at the UHUK conference in Birmingham in October, and was particularly well received. Looking to the future it would be ideal to disseminate this piece of work into the healthcare community for the benefit of out-of-hours providers across England, and therefore the patients. Finally, we suggested a solution for non-specialist users to access R functions on Excel spreadsheets through the software RExcel. Initially, this will be fine, but as discussed, care needs to be taken on the upkeep of these worksheets, as well as User training to avoid the 'garbage in - garbage out' effect.

In the uncertain times ahead, it is important for GP out-of-hours providers to have a good relationship with their commissioners to ensure renewal of contracts. Implementing the *CPD* and *Daily Profile* models will allow an organisation to optimise staffing. Monitoring GP triage lengths will allow them to improve performance, and both of these things will ensure the success of the out-of-hours GP providers.

Bibliography

- [1] UHUK Website. URL <http://www.urgenthealthuk.com/>.
- [2] Deadly new flu virus in US and Mexico may go pandemic. New Scientist, May 2009.
- [3] First uk swine flu patient death. BBC, June 2009. URL <http://news.bbc.co.uk/1/hi/uk/8099832.stm>.
- [4] Press releases and media updates on Swine Influenza. Health Protection Agency website, October 2010. URL <http://www.hpa.org.uk>.
- [5] KTP website, September 2010. URL <http://www.ktponline.org.uk>.
- [6] Equity and excellence: Liberating the NHS, July 2010. Crown Copyright 2010.
- [7] A. Deslauriers A. N. Avramidis and P L'Ecuyer. Modelling daily arrivals to a telephone call center. *Management Science*, 50(7):896–908, 2004.
- [8] T. Baier and E. Neuwirth. Excel :: Com :: R. *Computational Statistics*, 22(1):91–108, 2007.
- [9] K. P. Burnham and D. R. Anderson. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, New York, 2002.
- [10] Edited by J. M. Chambers and T. J. Hastie. *Statistical models in S*. Chapman Hall, London, 1993.
- [11] W. S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836, 1979.
- [12] W. S. Cleveland and S. J. Devlin. Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610, 1988.
- [13] Achim Zeileis David Meyer and Kurt Hornik. *vcd: Visualizing Categorical Data.*, 2010. R package version 1.2-9.
- [14] M. H. DeGroot and M. J. Schervish. *Probability and Statistics*. Addison-Wesley, New York, 2002.
- [15] N. R. Draper and H. Smith. *Applied Regression Analysis*. John Wiley & Sons, New York, 1998.
- [16] Paul James. North on alert as first swine flu case revealed. The Journal, May 2009.
- [17] D. Ladiray and B. Quenneville. *Seasonal Adjustment with the X-11 Method (Lecture Notes in Statistics; 158)*. Springer, New York, 2001.
- [18] C. Mallows. Some comments on C_p . *Technometrics*, 15(4):661–675, 1973.

- [19] Erich Neuwirth, with contributions by Richard Heiberger, Christian Ritter, Jan Karel Pieterse, and Jurgen Volkering. *RExcelInstaller: Integration of R and Excel, (use R in Excel, read/write XLS files)*, 2010. URL <http://CRAN.R-project.org/package=RExcelInstaller>. R package version 3.0-19.
- [20] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. URL <http://www.R-project.org>. ISBN 3-900051-07-0.
- [21] H. Shen and J. Z. Huang. Interday forecasting and intraday updating of call center arrivals. *Service Operations Management*, 10(3):391–410, 2008.